CONSTRUCTION AND USE OF CHIMERIC α-AMYLASES TO STUDY PROTEIN SECRETION IN *BACILLUS SUBTILIS*

KEITH STEPHENSON

Thesis submitted to the University of Newcastle upon Tyne for the degree of Doctor of Philosophy

September 1996

Acknowledgements

I would like to thank Colin Harwood for his supervision and guidance throughout this project and also for allowing me to achieve my potential. This work was funded partly by Novo Nordisk and I would like to thank all the members of the Bacterial Gene Technology Department (Bagsværd, Denmark) for making my stay a good one and in particular, Steen T Jørgensen, Christina L Jensen and Per Linå Jørgensen. I would like to express my sincere gratitude to all my family, but especially to my mother, for all the support over the years. Finally, thanks to Michelle for being there throughout the peaks and troughs of the life of a research student.

CONTENTS

List of figures

List of tables:

CHAPTER 1

SUMMARY

To investigate the influence of exoprotein charge on protein secretion, genes encoding a range of chimeric α-amylases with altered net charge were constructed by a PCR-based technique and expressed in *B. subtilis*. The chimeric α-amylases were based on AmyL and contained specific regions from two related *Bacillus* α-amylases, AmyQ and AmyS. The engineered changes were introduced into *amyL* to increase the net positive charge of the chimeric enzymes, when compared to wild type AmyL.

Isoelectric focusing confirmed that chimeric α-amylases possessed considerable positive charge and the temperature and pH optima of the most basic chimeric protein, AmyLQS50.5, were largely unaffected by the engineered changes. However, the structural stability, thermostability and the specific activity of this chimeric α-amylase were adversely affected.

In general, lower amounts of chimeric α-amylases were released into the culture supernatants. Pulse-chase experiments revealed that the rate of processing of AmyLQS50.5 was reduced when compared to AmyL and also that the mature forms of both α-amylases were subjected to degradation during or shortly after translocation, although the chimeric enzyme was affected most. When compared to AmyL, the rate of refolding of AmyLQS50.5 was reduced approximately 3-fold, maintaining this protein in a protease-sensitive conformation for an increased period of time. Therefore, it is proposed that the extensive co- or post-translocational degradation of the chimeric enzyme was a consequence of reduced folding kinetics on the outer surface of the cytoplasmic membrane since, when in its native conformation, AmyLQS50.5 is highly resistant to the activity of *B. subtilis* extracellular proteases. These observations have important implications for the use of *B. subtilis* as a host for the secretion of heterologous proteins.

The most positively charged chimera, AmyLQS50.5, was shown to bind significantly to cell walls isolated from *B. subtilis* 168, whereas AmyL and human serum albumin did not. This suggests that protein charge can influence the degree to which exoproteins interact with, and bind to, the cell wall as a consequence of electrostatic interactions with the anionic polymers.

CHAPTER 2

INTRODUCTION

2.1 Protein secretion in prokaryotes - introduction

All cells, eukaryotic or prokaryotic, possess the ability to translocate proteins across hydrophobic membranes and into cellular compartments other than the ones in which they were synthesized. In this study, unless indicated otherwise, the terms export and secretion will be used to refer to the process by which proteins are translocated across the cytoplasmic membrane. The terms are used irrespective of the nature or the final destination of the protein of interest.

Proteins destined for export (exoproteins) are synthesized as precursors which are targeted to the cytoplasmic membrane by a N-terminally located signal peptide. Exoproteins are held in a translocation-competent conformation by the action of soluble cytoplasmic proteins called chaperones and this interaction appears to be essential for export. Once the exoprotein has been targeted to the translocation site in the membrane, translocation occurs by an as yet undefined mechanism which is energy requiring. During translocation the signal peptide is removed from the preprotein by a membrane-associated peptidase and the mature exoprotein is released.

Bacteria of the Gram-positive genus *Bacillus* have a high capacity to secrete proteins directly into the culture supernatant (Priest, 1977). Coupled with their "generally regarded as safe" (GRAS) status, organisms of this genus have been used since the early 1900's for the production of enzymes for the food and detergent industries. The liberation of the proteins of interest directly into the culture supernatant allows subsequent purification and downstream processing to be carried out with relative ease.

The mechanism of protein secretion has been extensively studied in *Escherichia coli* and the components of the general secretory pathway (GSP) of this organism have been well characterized (2.2). Until relatively recently the mechanism of protein secretion in *Bacillus* species has received relatively little attention. However, with the development of robust molecular biological techniques for the manipulation of *B. subtilis*, knowledge of the mechanisms of protein secretion in this model Gram-positive organism has expanded considerably in recent years. Elucidation of the components required to export proteins from the *B. subtilis* cell will not only add to the growing knowledge of the molecular biology and physiology of this organism, but is likely to provide a means of optimizing the range of proteins exported and also their yields.

2.2 Protein secretion in *E. coli*

The GSP of *E. coli* has been comprehensively reviewed (Pugsley, 1993) and to introduce the key components of the protein export machinery, the GSP

will be described briefly in this section. Some *E. coli* exoproteins are synthesized without signal peptides and are exported by non-GSP accessory pathways. These exoproteins will not be considered further.

2.2.1 Chaperones

Following translation, the nascent preprotein emerges from the ribosome and unfolded parts of the exoprotein are recognized and bound by soluble cytoplasmic proteins known as chaperones. The binding of chaperone molecules to the exoprotein precursor arrests folding and maintains the exoprotein in a conformation which is competent for translocation across the cytoplasmic membrane. The product of the *secB* gene appears to be the major GSP exoprotein chaperone in *E. coli* (Kumamoto *et al.*, 1989; Kumamoto, 1989; Kumamoto, 1991). Other chaperones have been identified in *E. coli* (such as DnaK, DnaJ, GroEL and GroES), but these proteins have additional functions in systems other than protein export (Phillips and Silhavy, 1990; Wild *et al.*, 1992). *E. coli* is still capable of translocating exoproteins in the absence of SecB and therefore, other chaperones must be capable of assuming the role of SecB.

Following chaperone binding, the preprotein/chaperone complex is targeted to export sites within the cytoplasmic membrane and evidence suggests that the signal peptide penetrates the lipid bilayer of the membrane (Killian *et al.*, 1990a; Killian *et al.*, 1990b). The targeting is achieved by the affinity of a peripheral membrane protein, SecA, for signal peptides, as yet undetermined regions of the mature protein and also SecB (Lill *et al.*, 1990; Hartl *et al.*, 1990).

2.2.2 The *E. coli* translocation complex

Initially the translocation complex, or translocase, of *E. coli* was thought to be composed of the peripheral membrane protein SecA and just two integral membrane proteins, SecY and SecE (Brundage *et al.*, 1990; Akimaru *et al.*, 1991), which form a stable complex in the cytoplasmic membrane (Bieker and Silhavy, 1990). However, additional studies have identified three other integral membrane components of the translocation complex, namely SecD, SecF and SecG (Gardel *et al.*, 1990; Brundage *et al.*, 1992; Nishiyama *et al.*, 1993; Nishiyama *et al.*, 1994).

The SecA protein is located on the inner side of the cytoplasmic membrane in association with the integral membrane components. SecA exists as a dimer which binds and hydrolyses ATP to drive the translocation process (Lill *et al.*, 1989; Geller, 1991). It is likely that the hydrolysis of ATP by SecA releases the preprotein from SecA which then crosses the cytoplasmic membrane

using the proton-motive force (Driessen, 1990). The integral membrane components of the translocation complex form a channel in the cytoplasmic membrane through which exoproteins pass during translocation. It has been proposed that this channel opens laterally in the membrane to enclose the exoprotein, the N-terminal of which has penetrated the lipid bilayer due to the signal peptide (Simon and Blobel, 1992).

2.2.3 Translocation and release from the cytoplasmic membrane

The exoprotein precursor protein moves through the channel created by the translocation complex and during, or shortly after translocation, the signal peptide is removed from the precursor by the action of another integral membrane protein, the signal peptidase. The main signal peptidase is the product of the *lep* gene and this protein is responsible for the processing of the majority of *E. coli* precursors and also the M13 coat protein precursor (Dalbey and Wickner, 1985). Following processing, the mature exoprotein is released from the cytoplasmic membrane and folds into its native conformation, possibly facilitated by SecD and SecF (Gardel *et al.*, 1990; Bieker-Brady and Silhavy, 1992). Also periplasmic chaperone molecules with disulphide bond isomerase (Kamitani *et al.*, 1992) and peptidylprolyl-*cis-trans*-isomerase activity (Liu and Walsh, 1990) have been implicated in these late-stage folding events. Following folding, terminal branches of the GSP transport specific exoproteins to their final locations, which may or may not be in the extracellular milieu.

2.3 The properties of secretory proteins

Exoproteins secreted using the GSP in *E. coli*, and those exported by equivalent Gram-positive systems, have certain features which distinguish them from intracellular proteins and facilitate their secretion.

2.3.1 Signal peptides

The primary sequence of an exoprotein must include a signal peptide for it to be secreted. Signal peptides are exclusively N-terminal and direct the translocation of exoproteins across the cytoplasmic membrane of prokaryotes or the endoplasmic reticulum (ER) membrane of eukaryotes. The signal peptide is the only currently recognized characteristic that differentiates exoproteins from cytoplasmic proteins.

Signal peptides exhibit three structural characteristics (Fig. 2.1); a positively charged N-terminal domain (N-region), a stretch of hydrophobic amino acid residues (H-region) and a more polar region (C-region) which contains the

consensus signal peptidase cleavage site. The cleavage site is typically Ala-X-Ala-↓-, where X represents any amino acid and ↓ represents the peptide bond which is cleaved by the signal peptidase (von Heijne, 1985).

The signal peptides of *Bacillus* species and other Gram-positive organisms are longer than their Gram-negative or eukaryotic counterparts. For example, the signal peptides of *Bacillus* exoproteins are commonly between 30 and 40 amino acid residues in length compared with an average of 22 residues in other organisms. The extra length observed in Gram-positive genera is distributed across the whole of the signal peptide and not restricted to one particular region. Also, the N-region of Gram-positive signal peptides tend to have a higher concentration of positively charged amino acids than the corresponding region from *E. coli* signal peptides (Simonen and Palva, 1993).

Bacterial exoproteins that become covalently attached to the external side of the cytoplasmic membrane or the outer membrane of Gram-negatives by lipid moieties are known as lipoproteins. The lipid group is covalently attached to an N-terminal cysteine residue of the mature protein (Wu and Hayashi, 1986). The signal peptides of lipoproteins are similar in structure to those of other secretory proteins with N-, H- and C-regions. However, the C-region is shorter and the consensus cleavage sequence is Leu-Ala-Gly-↓-Cys (von Heijne, 1989).

2.3.2 Propeptides

Exoproteins of *Bacillus* species often have additional amino acid residues located between the signal peptide and the mature regions. These sequences, called propeptides, are removed following the cleavage of the signal peptide and can vary in size from 8 amino acids *e.g. B. subtilis* α-amylase (Takase *et al.*, 1988) up to 204 amino acid residues *e.g. Bacillus stearothermophilus* neutral protease (Takagi *et al.*, 1985). Although several roles have been proposed for propeptides, it is thought that their main role in secretion involves the folding and maturation of exoproteins at a stage following translocation (Cash *et al.*, 1989; Zhu *et al.*, 1989; Simonen and Palva, 1993).

2.4 The *B. subtilis* export machinery

In *E. coli*, translocated exoproteins can have one of several final locations, depending on the nature of a particular protein. Exoproteins can remain associated with the cytoplasmic or outer membranes if they have the required structural characteristics. Conversely, following translocation, exoproteins may exist free in the periplasm or cross the outer membrane and become truly extracellular. *B. subtilis* has a much simpler envelope structure than *E. coli*, with

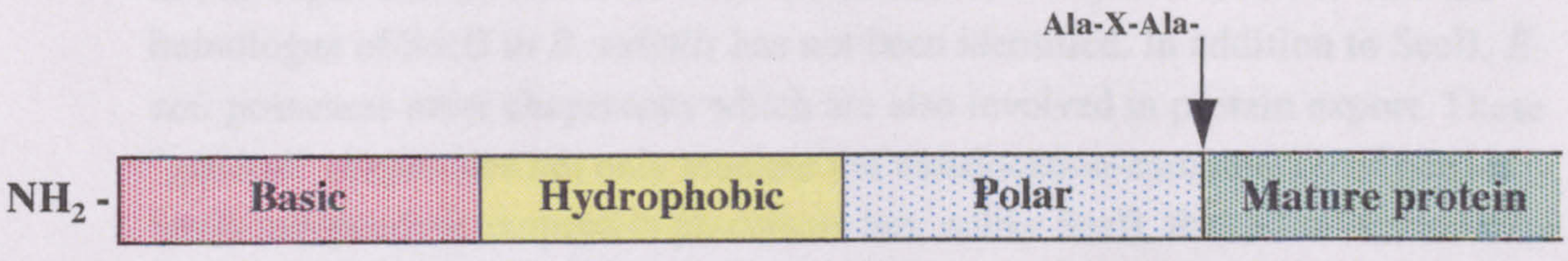


Fig. 2.1. The basic structure of signal peptides.
The arrow indicates the signal peptidase cleavage site.

only a thick cell wall lying external to the cytoplasmic membrane (2.6). Therefore, translocated proteins can remain associated with the cytoplasmic membrane or the cell wall, or move through the wall and reside in the extracellular milieu. As with *E. coli*, the final destination of *B. subtilis* exoproteins is dependent upon the nature of the protein concerned. This section will provide a broad overview of the identified components of the *B. subtilis* export pathway some, but not all of which, are homologous to components of the *E. coli* GSP.

2.4.1 Chaperones

In any export system chaperones are required to maintain exoprotein precursors in translocation-competent conformations. The chaperones identified to date in *E. coli* have been discussed (2.2.1) and the main exoprotein chaperone in this organism appears to be SecB (Kumamoto, 1991). However, to date the homologue of SecB in *B. subtilis* has not been identified. In addition to SecB, *E. coli* possesses other chaperones which are also involved in protein export. These "general" chaperones not only mediate the translocation-competency of certain SecB-independent exoprotein precursors but, unlike SecB, they also facilitate the proper folding of intracellular proteins (Kumamoto, 1991). In *E. coli* at least two chaperone systems appear to operate in secretion. The first is dependent on SecB and is specific for proteins targeted to the export pathway. The second pathway is SecB-independent and is mediated by a more general set of chaperones which affect the folding of intracellular, as well as, extracellular proteins.

When cells are subjected to certain stresses, such as heat shock, the likelihood of protein denaturation and aggregation is increased. Chaperones help to reduce protein aggregation and denaturation and, as a consequence, many chaperones have been classified as stress or heat shock proteins (HSP's). In *E. coli* some of the chaperones involved in the cells response to heat shock are encoded by the *groESL* and *dnaK* operons. *E. coli* temperature sensitive *groESL* mutants were shown to be defective in the export of β-lactamase (Kusukawa *et al.*, 1989) and purified GroEL was shown to bind to certain exoprotein precursors and to maintain their translocation competence (Lecker *et al.*, 1989). Further, the export of a LamB-β-galactosidase hybrid protein was promoted by the simultaneous overexpression of the *dnaK* gene (Phillips and Silhavy, 1990). Since in *E. coli* some of the chaperones encoded by heat shock genes are involved in protein export, it is possible that the homologous chaperones from *B. subtilis* also play a role in secretion.

2.4.1a The groESL *operon*

In *B. subtilis* there are at least three classes of heat shock genes. Those making up class I, such as the genes of the *groESL* and *dnaK* operons, are expressed from a vegetative promoter (σ^A) (Chang *et al.*, 1994). Class I genes are only induced by heat stress whereas those of the other classes are inducible by other stresses such as osmotic shock or glucose limitation.

As in *E. coli*, the *B. subtilis groEL* and *groES* genes are organized in an operon with the gene order *groES-groEL* and are transcribed from a promoter upstream of *groES* (Fig. 2.2) (Schmidt *et al.*, 1992; Li and Wong, 1992). The promoter is utilized under normal conditions but, following heat shock, transcription of the operon is induced up to 18-fold. Immediately downstream from the transcription start site is a nine base pair (bp) inverted repeat (IR) which is separated by a nine bp spacer region. A similar IR has also been found upstream of the *B. subtilis dnaK* operon (Wetzstein *et al.*, 1992) and the *dnaK* and *groESL* operons *of Clostridium acetobutylicum* (Narberhaus and Bahl, 1992; Narberhaus *et al.*, 1992). It has been postulated that this IR has a role in the regulation of the heat shock response by preventing high level expression of the operon under normal conditions (2.4.1c).

In *E. coli* GroEL is a tetradecameric protein composed of two stacked rings. The individual rings contain seven subunits and each subunit has a molecular mass of approximately 60 kDa. Together the rings form a cylinder with a central cavity which is able to bind 1 to 2 molecules of unfolded protein. GroES is a single heptameric ring of 10 kDa subunits which forms a 1:1 complex with GroEL by binding to one end of the cylinder.

The *groEL* and *groES* genes from *B. subtilis* have been cloned, encode proteins of 58 and 10 kDa, respectively, and are essential for growth and viability under normal and heat shock conditions (Schmidt *et al.*, 1992; Li and Wong, 1992). At the amino acid level, GroEL and GroES from *B. subtilis* have been shown to be 78.7% and 54.9% identical, respectively, to the corresponding proteins from *E. coli* (Schmidt *et al.*, 1992). Electron microscopy has shown the GroEL particle from *B. subtilis* to have a similar morphology to that from *E. coli* (Carrascosa *et al.*, 1990). Additionally, the *B. subtilis* GroEL homologue cross-reacts with antiserum raised against *E. coli* GroEL showing that the proteins from the different bacteria must have at least one antigenic determinant in common.

2.4.1b The dnaK *operon*

The *dnaK* operon of *B. subtilis* has been cloned, encodes at least four genes involved in the heat shock response and is organized as follows; *hrcA*-

grpE-dnaK-dnaJ (Fig. 2.2) (Wetzstein *et al.*, 1992). The first gene in the operon, *hrcA*, encodes a polypeptide with a molecular mass of 39 kDa which is a negative regulator involved in controlling the operon (Yuan and Wong, 1995; Schulz and Schumann, 1996).

The *B. subtilis grpE* gene codes for a protein with a predicted molecular mass of 22 kDa. Sequence alignments of GrpE proteins from different species revealed only a low level of homology between the GrpE proteins. For example, GrpE from *E. coli* and *B. subtilis* exhibit only 30.5% amino identity. However, within the GrpE proteins there are two regions of increased homology. These regions, termed regions I and II, are 65% and 52.4% homologous, respectively, when compared with the corresponding regions from *E. coli* GrpE (Wetzstein *et al.*, 1992). It is possible that the homologous regions in these protein are involved in the interaction of GrpE with other proteins such as DnaK since these two proteins are known to interact in *E. coli* (Johnson *et al.*, 1989).

The *dnaK* gene encodes a protein with a predicted molecular mass of 66 kDa which has a good overall homology with other DnaK proteins. However, homology decreases towards the C-termini of the proteins. The *B. subtilis* DnaK protein has 59.4% overall amino acid identity to *E. coli* DnaK. Wetzsein *et al* (1992) analyzed the DnaK proteins from a number of different organisms and found that, when compared to Gram-negative organisms, the DnaK proteins from *B. subtilis* and other Gram-positives had a deletion of 24 amino acid residues near the N-termini. The missing region corresponded to *E. coli* DnaK amino acid residues 78 to 102 and may represent a unique and distinguishing feature of the DnaK proteins from Gram-positive origin.

The three-dimensional crystal structure of the N-terminal region of a bovine heat shock protein, HSC70, has been reported (Flaherty *et al.*, 1990). Interestingly, when this protein is compared with *E. coli* DnaK, residues 78 to 102 correspond to a region which is located on the outside of the protein and forms an appendage with no known biological function. This suggests that this region is not required for biological activity of the chaperone and may explain why this region has not been retained by the Gram-positive organisms studied.

dnaJ is the last gene in the *B. subtilis dnaK* operon and the predicted molecular mass of the encoded protein is 41 kDa. The *dnaJ* gene has not been studied as extensively in *B. subtilis* as the other genes in the operon. In *E. coli* the DnaJ protein, in co-operation with DnaK and GrpE, stimulates the recycling of DnaK after its interaction with precursor proteins (Liberek *et al.*, 1991). It is possible that DnaJ has a similar function in *B. subtilis*.

2.4.1c Regulation of the groESL *and* dnaK *operons*

The *groESL* and *dnaK* operons are controlled at the level of transcription and the amount of operon-specific mRNA can be seen to increase transiently following heat shock (Schmidt *et al.*, 1992; Wetzstein *et al.*, 1992; Schulz *et al.*, 1995).

Inspection of these operons reveals the presence of putative vegetative promoter sequences upstream from the transcription initiation sites which are activated post heat shock (Schmidt *et al.*, 1992; Wetzstein *et al.*, 1992). Detailed analysis of the sequences near the transcription start sites revealed the presence of 9 bp IR's separated by a 9 bp spacer region (5'-TTAGCACTC-N_9-GAGTGCTAA-3'). Further, the same IR can also be found upstream of the *groESL* and *dnaK* operons of numerous Gram-positive and Gram-negative bacterial species. Since these genes encode proteins that act as molecular chaperones, the IR has been termed CIRCE, the controlling IR of chaperone expression (Zuber and Schumann, 1994).

The exact details of the mechanism by which CIRCE controls the expression of the *dnaK* and *groESL* operons has yet to be elucidated. However, mutational studies on the *dnaK* operon have shown that CIRCE is a negative *cis* acting controlling element and that its activity is independent of the promoter responsible for the transcription of the operon (Zuber and Schumann, 1994). It is likely that CIRCE forms a secondary structure in either the DNA or RNA to which a repressor protein, possibly HrcA, binds under non-heat shock conditions allowing only a basal level of transcription. Presumably the putative repressor protein would be unable to function under heat shock conditions.

The *B. subtilis* chaperones identified to date are all involved in the cellular response to heat shock. It is likely that this organism possesses other, non-heat shock response chaperones, for example, a counterpart to the major *E. coli* exoprotein chaperone, SecB.

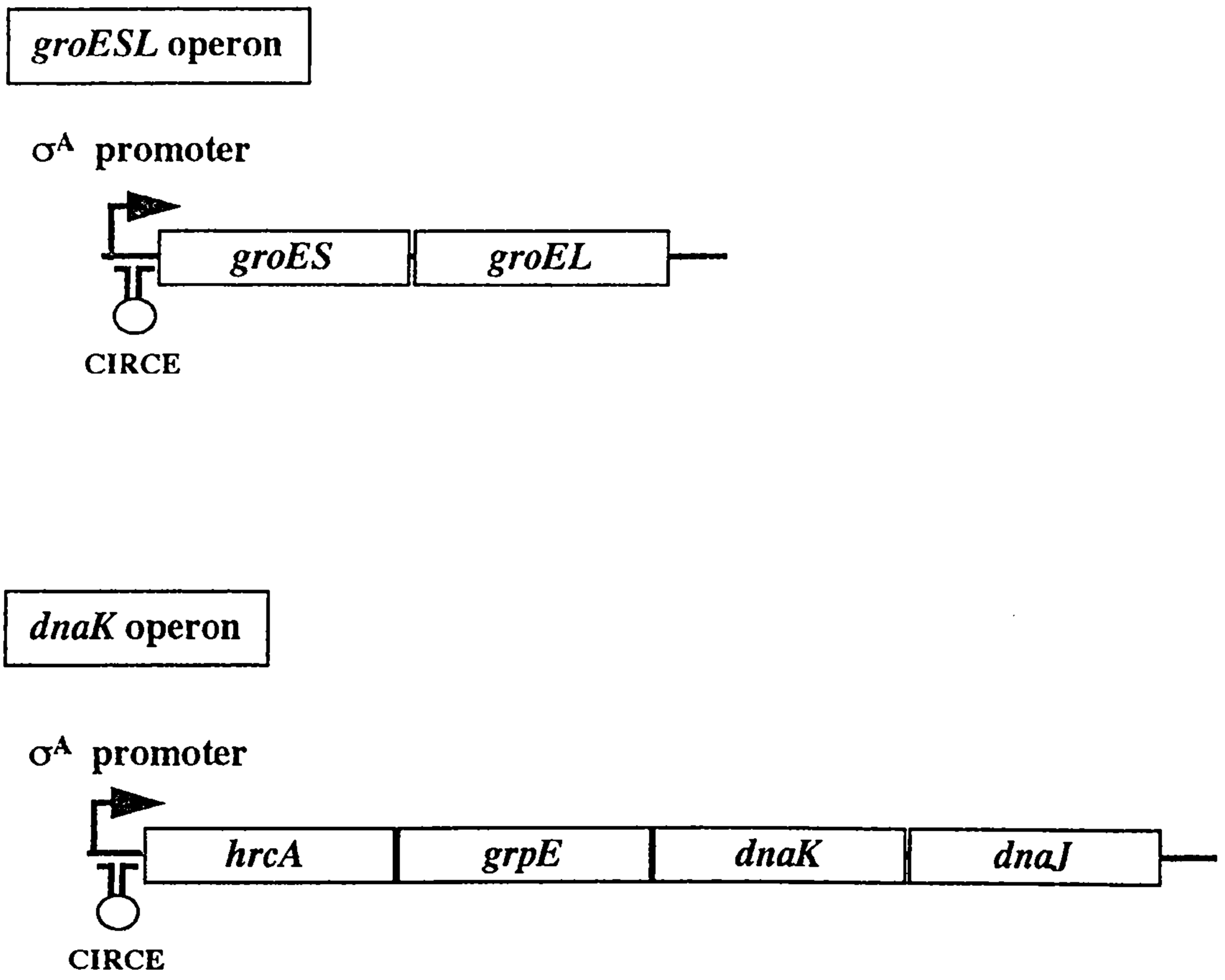


Fig. 2.2. The organization of the *B. subtilis groESL* and *dnaK* operons.

2.4.2 The *Bacillus subtilis* translocation complex

Following chaperone binding, the preprotein is presented to the translocation complex in the cytoplasmic membrane and the identified components of this complex are discussed below and shown schematically in Fig. 2.3.

2.4.2a SecA protein

In *E. coli* the peripheral membrane protein SecA has a key role in secretion by coupling ATP hydrolysis to the translocation of the preprotein across the cytoplasmic membrane (Wickner *et al.*, 1991). The *B. subtilis* homologue of the *E. coli secA* gene has been identified, cloned and sequenced (Miyakawa and Komano, 1980; Sadaie *et al.*, 1991). The *B. subtilis* SecA protein is the product of the *div* gene and was identified by a mutation in a gene controlling the initiation of septation and has been subsequently shown to be required for various post-exponential phase phenomena (Sadaie and Kada, 1983; Sadaie and Kada, 1985). Sequence analysis revealed 47.2% and 50% nucleotide and amino acid identity, respectively, when compared to the corresponding gene or protein from *E. coli* and even greater homologies can be detected in certain regions of the proteins (Sadaie *et al.*, 1991).

A *B. subtilis secA* temperature sensitive mutant strain, *div-341*^{ts}, was isolated which accumulated α-amylase and β-lactamase precursors and exhibited temperature sensitive secretion of α-amylase and protease (Sadaie, 1989; Sadaie and Kada, 1985). The SecA protein of *B. subtilis* is a homodimer of 94 kDa subunits which cross-reacts weakly with antiserum raised against *E. coli* SecA (Takamatsu *et al.*, 1992). During translocation, ATP is hydrolyzed by SecA to drive the precursor across the translocation complex (Driessen, 1990; Wickner *et al.*, 1991). The ATPase activity of purified *B. subtilis* SecA was investigated and was found to be three times higher than that of the *E. coli* protein, assayed under the same conditions (Takamatsu *et al.*, 1992).

Takamatsu *et al* (1992) also demonstrated that *B. subtilis* SecA could not promote the translocation of an OmpF-Lpp hybrid protein in an *in vitro* translocation assay, unlike *E. coli* SecA. Further, the *E. coli* and *B. subtilis secA* genes were unable to complement *secA* mutations *in vivo* in *B. subtilis* or *E. coli*, respectively. Therefore, on the basis of nucleotide and amino acid similarity, it seems as though the SecA proteins from these two organisms are related but are not sufficiently alike to be functionally complementary.

2.4.2b SecY protein

In *E. coli* the product of the *secY* (*prlA*) gene, SecY, is a 49 kDa integral membrane protein which is essential for growth and the translocation of exoproteins (Shiba *et al.*, 1984; Akiyama and Iko, 1987). SecY is a component of the translocator through which exoproteins pass during translocation across the cytoplasmic membrane. The *E. coli* SecY protein is a transmembrane protein with 10 membrane spanning regions and 11 hydrophilic regions, 6 of which are exposed to the cytoplasm and 5 to the periplasm (Akiyama and Iko, 1987; Watanabe and Blobel, 1989).

The *B. subtilis* homologue of *secY* has been cloned and sequenced and codes for a 47 kDa protein which has 41% overall amino acid identity with the corresponding protein from *E. coli* (Suh *et al.*, 1990; Nakamura *et al.*, 1990). Analysis of the amino acid sequence of *B. subtilis* SecY revealed that, as in *E. coli*, the protein had 10 potential membrane spanning regions which were 20-30 amino acids long and rich in hydrophobic residues (Suh *et al.*, 1990; Nakamura *et al.*, 1990). Further, the hydrophobicity plots of the *E. coli* and *B. subtilis* SecY proteins were similar and the charge distribution of *B. subtilis* SecY was typical of transmembrane proteins, indicating that the protein from *B. subtilis* also resides in a transmembrane location.

A more detailed analysis of the amino acid sequences of the *B. subtilis* and *E. coli* SecY proteins revealed higher levels of homology within specific regions. Cytoplasmic regions II, III, IV and V exhibited from 44% to 62% identity. The regions with the highest degree of homology are the putative membrane spanning regions. For example, transmembrane regions I, II, IV, V, VII and X shared between 50% and 73% amino acid identity, although the latter homology may simply reflect the limited range of amino acids capable of taking up a transmembrane location rather than highly conserved regions of the SecY proteins.

Nakamura *et al* (1990) proposed that the conserved regions of the complete proteins were essential for the proper functioning of the SecY proteins and that those exposed to the cytoplasm may interact with common components of the secretion machinery and/or the signal peptides of exoprotein precursors. In contrast, cytoplasmic regions I and VI, which represent the N- and C-terminal regions, respectively, exhibited almost no homology, as did the regions which correspond to the periplasmic domains of *E. coli* SecY.

Complementation studies showed that the *B. subtilis secY* gene was unable to restore the growth of an *E. coli secY* mutant (*secY24*) and that the same gene was detrimental to the growth of an *E. coli* strain containing its wild type *secY*

gene (Suh *et al.*, 1990). Additionally, the presence of the *B. subtilis secY* gene was shown to reduce the secretion of the maltose binding protein and OmpA in *E. coli secY24* when compared to the wild type strain. These observations suggest that the *B. subtilis* SecY protein interferes directly or indirectly with the protein export machinery of *E. coli* and Suh *et al* (1990) postulated that the above observations may be caused by a titration effect of the overexpressed *B. subtilis* SecY protein with an essential *E. coli* secretion factor.

2.4.2c SecE protein

E. coli SecE is associated with SecY in the cytoplasmic membrane and together constitute the translocator *via* which exoproteins pass during translocation. The *E. coli* SecE protein is a 127 amino acid polypeptide with a molecular mass of 13.6 kDa and Schatz *et al* (1989) have suggested that it is a transmembrane protein with 3 membrane-spanning domains. The N- and C-termini of *E. coli* SecE are believed to be located on the cytoplasmic and periplasmic sides of the cytoplasmic membrane, respectively. Initially, the complete SecE protein was thought to be essential for protein translocation and cell viability (Schatz *et al.*, 1989) but deletion studies suggested that only one transmembrane domain in the C-terminus is required for the activity of the *E. coli* SecE protein (Schatz *et al.*, 1991).

Recently the *B. subtilis* homologue of the *E. coli secE* gene has been cloned, sequenced and analysed (Jeong *et al.*, 1993). The *B. subtilis secE* gene codes for a protein of 59 amino acids which shares just 21% amino acid identity with the C-terminal region of *E. coli* SecE. Further, the *B. subtilis* SecE protein has a putative transmembrane domain and a similar hydrophobicity plot to the C-terminus of *E. coli* SecE. The evidence suggests that the functionally active domain of *E. coli* SecE resides in the C-terminus of the protein (Schatz *et al.*, 1991). This region shows some homology to the *B. subtilis* SecE protein, suggesting that the functional domain of the SecE proteins have been conserved.

Complementation studies were carried out using the *B. subtilis secE* gene and a cold-sensitive *E. coli secE* mutant (*secEcsE501*) (Jeong *et al.*, 1993). Using Western blotting and pulse-chase techniques, the *B. subtilis secE* gene was able to restore growth and translocation defects of the *E. coli secE* mutant at restrictive temperatures. Jeong *et al* (1993) have also putatively identified the gene coding for the *Bacillus licheniformis* SecE homologue which shows 83.3% amino acid identity to *B. subtilis* SecE. Initial analysis of the putative *B. licheniformis* SecE protein indicated that, like *B. subtilis* SecE but in contrast to *E. coli* SecE, the protein has only one transmembrane domain.

2.4.2d Summary

The components of the *B. subtilis* translocation complex identified to date have been discussed above and it is likely that others will be identified in the future. The identified components exhibit varying degrees of similarity when compared with the corresponding proteins from the well characterized *E. coli* GSP. However, at the functional level most of the components are relatively well conserved. Some components of the *B. subtilis* translocation complex are able to complement defects in the homologous proteins from *E. coli* and others are not. This indicates that the targeting and translocation complexes have been relatively well conserved throughout evolution but individual components have been co-adapted to specific systems which prevents them from being completely functionally interchangeable. This reflects the major differences between the evolutionary distinct Gram-positive and Gram-negative bacteria and may be a consequence of their different cell envelope architectures.

2.4.3 *B. subtilis* Ffh - a homologue of the eukaryotic SRP54?

In eukaryotic cells, targeting of secretory proteins to the ER is mediated by a ribonucleoprotein particle, termed the signal recognition particle (SRP). The SRP is composed of a single molecule of RNA (SRP 7S RNA) and six polypeptides of varying molecular masses (Walter and Blobel, 1982; Siegel and Walter, 1988). The polypeptide subunit responsible for binding to the signal peptide of the nascent polypeptide has a molecular mass of 54 kDa and is therefore termed SRP54 (Krieg *et al.*, 1986; Kurzchalia *et al.*, 1986; Wiedmann *et al.*, 1987).

SRP binds to the emerging signal peptide of the nascent polypeptide and blocks its translation. SRP then facilitates the binding of the SRP/ribosome/nascent polypeptide complex to a receptor, known as the docking protein, on the ER (Meyer *et al.*, 1982). Translation of the polypeptide is resumed after the dissociation of SRP and the polypeptide is then translocated across the membrane of the ER cotranslationally. There is some evidence which suggests that an SRP-like export pathway exists in *E. coli* (Ribes *et al.*, 1990; Luirink *et al.*, 1992; Phillips and Silhavy, 1992). In this organism the 48 kDa product of the *ffh* gene has been identified as the homologue of the eukaryotic SRP54 protein and the *ftsY* gene product showed some homology to the α subunit of the docking protein (Romisch *et al.*, 1989).

Recently, the *E. coli ffh* gene has been used as a probe to identify the *B. subtilis ffh* gene (Honda *et al.*, 1993). Subsequent analysis revealed that the *B. subtilis* Ffh protein shared 53.9% and 32.6% amino acid identity with *E. coli* Ffh

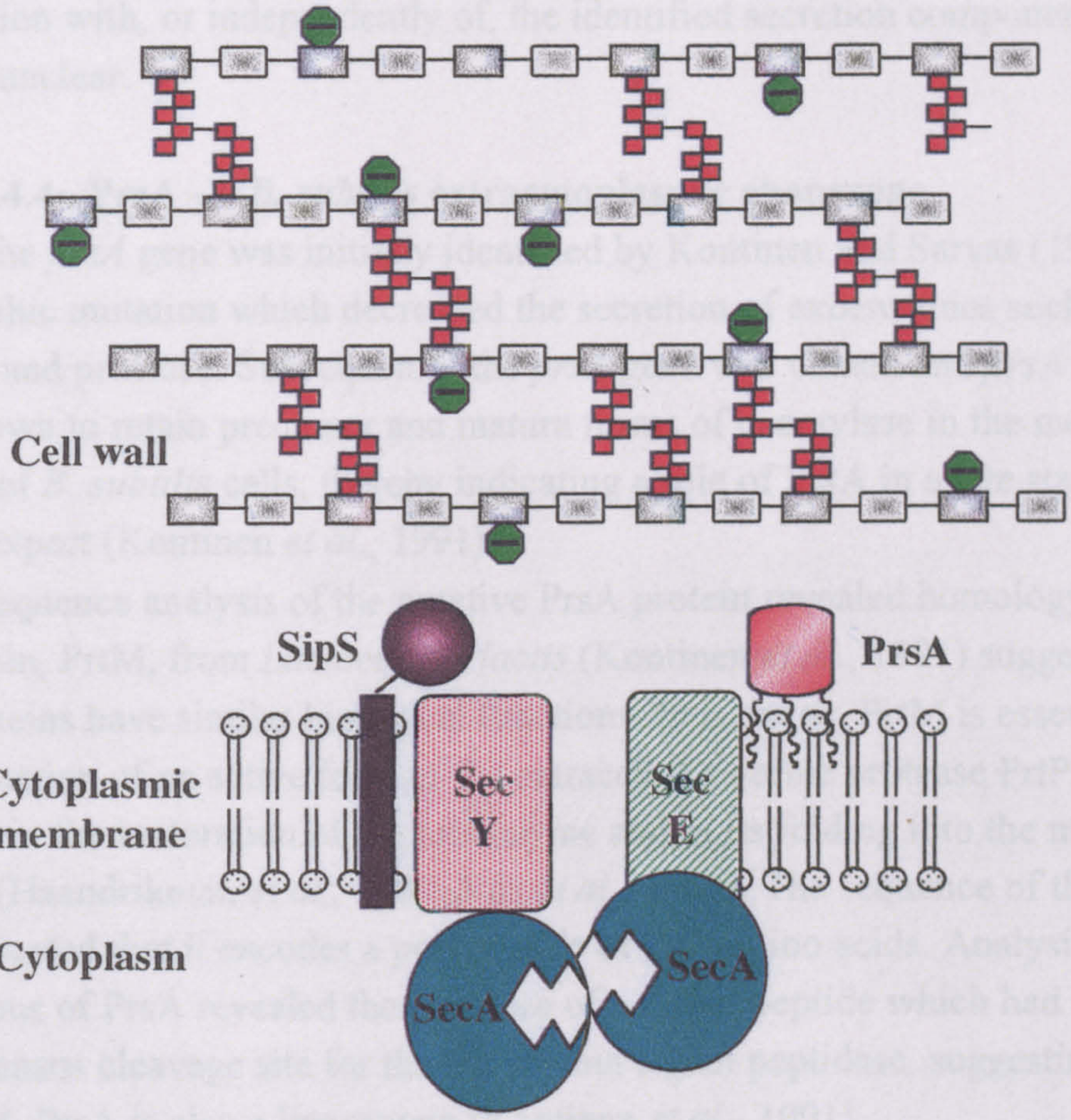
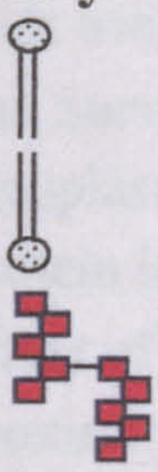


Fig. 2.3. Schematic representation of the identified components of the *B. subtilis* export machinery and the cell wall.

Key:



Lipid bilayer

Glycan chain

Cross-linking peptide

Anionic polymer

and canine SRP54, respectively. Therefore, in addition to the exoprotein targeting functions of already identified components of the secretion machinery, *e.g.* SecA in *B. subtilis* and *E. coli*, and SecB in *E. coli* only, it is possible that there are other cytoplasmic components which contribute to the targeting of the preprotein to the cytoplasmic membrane. However, whether an SRP-like system operates in conjunction with, or independently of, the identified secretion components remains unclear.

2.4.4 PrsA - a *B. subtilis* extracytoplasmic chaperone

The *prsA* gene was initially identified by Kontinen and Sarvas (1988) as a pleiotrophic mutation which decreased the secretion of exoenzymes such as α-amylase and protease. Subsequently the *prsA* gene was cloned and *prsA* mutants were shown to retain precursor and mature forms of α-amylase in the membrane fraction of *B. subtilis* cells, thereby indicating a role of PrsA in a late stage of protein export (Kontinen *et al.*, 1991).

Sequence analysis of the putative PrsA protein revealed homology with a lipoprotein, PrtM, from *Lactococcus lactis* (Kontinen *et al.*, 1991) suggesting that both proteins have similar biological functions. In *L. lactis*, PrtM is essential for the production of an active form of the extracellular serine protease PrtP, and is involved in the maturation of the proenzyme and in its folding into the mature enzyme (Haandrikman *et al.*, 1989; Vos *et al.*, 1989). The sequence of the *prsA* gene indicated that it encodes a polypeptide of 292 amino acids. Analysis of the N-terminus of PrsA revealed the presence of a signal peptide which had within it the consensus cleavage site for the lipoprotein signal peptidase, suggesting that, like PrtM, PrsA is also a lipoprotein (Kontinen *et al.*, 1991).

Since most *Bacillus* lipoproteins exist in a membrane-anchored form, Kontinen and Sarvas (1993) investigated the cellular location of the PrsA protein. They found that *E. coli* and *B. subtilis* strains containing a plasmid-encoded *prsA* gene had an abundant 33 kDa protein present exclusively in the membrane fraction of cells. Immunoblotting with anti-PrsA antiserum confirmed that this 33 kDa protein was the product of the *prsA* gene. Further, labelling experiments with tritiated fatty acids also localized PrsA to the membrane fraction (Kontinen and Sarvas, 1993). Therefore, the evidence suggests that PrsA is linked to the cytoplasmic membrane *via* the N-terminal lipid moiety with the remainder of the protein located on the outer side of the cytoplasmic membrane. This has the effect of concentrating PrsA at the site where it appears to function, *i.e.* in close proximity to the exoprotein as it emerges from the translocation complex.

In *E. coli* none of the identified components of the export machinery are located exclusively on the periplasmic side of the cytoplasmic membrane. However, the integral membrane SecF and SecD proteins do have large periplasmic domains (Gardel *et al.*, 1990). No sequence homology could be detected between the periplasmic domains of either SecF or SecD and PrsA, indicating that PrsA is a novel component of the secretion machinery of *B. subtilis* and probably other Gram-positive bacterial species (Kontinen *et al.*, 1991).

The role of PrsA in *B. subtilis* protein secretion was investigated using strains which overproduced *Bacillus* degradative extracellular enzymes (Kontinen and Sarvas, 1993). Simultaneous overexpression of PrsA and α-amylase increased the yield of α-amylase released into the supernatant by up to 6-fold. In contrast, overexpression of PrsA did not increase the secretion of native *B. subtilis* exoenzymes that were expressed naturally at a lower level, suggesting that PrsA is only able to enhance the secretion of overexpressed enzymes.

The viability of strains with an interrupted chromosomal copy of the *prsA* gene is influenced by the level of secretion of exoenzymes (Kontinen and Sarvas, 1993). When the chromosomal *prsA* gene was insertionally inactivated in a strain overexpressing α-amylase, the resulting transformants were not viable. However, when the same *prsA* inactivation was carried out on a strain devoid of the overexpressed α-amylase, viable cells were produced initially, which either grew with a disturbed colony morphology or which lost viability completely after a few generations.

Initial evidence suggested that PrsA functioned late in the protein export process (Kontinen *et al.*, 1991) and this was further supported by pulse-chase and immunoprecipitation studies using subtilisin C-PhoA hybrids in wild type and *prsA3* mutant backgrounds (Jacobs *et al.*, 1993). In the wild type strain the hybrid protein was processed from the precursor to the pro-form and then to the mature protein as expected. However, in the *prsA3* mutant strain several smaller proteins were precipitated using anti-PhoA antiserum, with only small amounts of the pre-, pro- or mature protein detectable for the duration of the experiment.

Additionally, the small amount of mature protein detectable lacked alkaline phosphatase activity. The smaller but related proteins are likely to be breakdown products of the SubC-PhoA protein. The pattern of degradation observed is indicative of increased sensitivity to endogenous proteases which in turn reflects a conformational change in the SubC-PhoA protein in the *prsA3* background. This, and the fact that the kinetics of signal peptide removal in the *prsA3* mutant and the wild type are essentially the same (Jacobs *et al.*, 1993),

confirm that PrsA acts after this processing step has occurred. In the absence of PrsA the secreted protein is unable to fold correctly and take up its final conformation. Consequently, the incorrectly folded protein is more susceptible to cleavage by proteases and is rapidly degraded.

In summary, the product of the *prsA* gene is a novel component of the *B. subtilis* secretion apparatus which is membrane anchored, located exclusively on the outer surface of the cytoplasmic membrane and is essential for protein export and normal cell growth. The current evidence suggests that PrsA acts as an extracytoplasmic chaperone which facilitates folding of exoproteins into their native conformations following translocation.

Although no obvious PrsA homologue has been identified in *E. coli*, this organism does possess periplasmic chaperone molecules which are involved in protein folding following translocation. One such protein, the product of the *dsbA* (*ppfA*) gene, is a disulphide isomerase and is necessary for the proper folding of certain exoproteins by mediating the formation of correct disulphide bonds (Kamitani *et al.*, 1992; Yamanaka *et al.*, 1994). Another type of *E. coli* periplasmic chaperones involved in protein folding are the peptidylprolyl-*cis*-*trans*-isomerases (PPI's) (Liu and Walsh, 1990). PPI's catalyze the isomerization of peptide bonds which have a proline residue in the second position and accelerate protein folding. The *B. subtilis* PrsA protein has been identified as a putative PPI (M Sarvas, personal communication) although this has not been demonstrated conclusively.

The evidence described above suggests that PrsA is a rate-limiting component of the export pathway which, when expressed at a high level, can increase the secretion of overexpressed exoenzymes several-fold. This provides a potential route for enhancing the secretion capacity of *B. subtilis* simply by increasing the expression of a single component of the export machinery.

2.4.5 Signal peptidases

As the exoprotein precursor is translocated across the cytoplasmic membrane it is processed into the mature form by the action of an integral membrane protein, the signal peptidase. Processing involves the enzymatic cleavage of the protein at a specific site to remove the signal peptide from the N-terminus of the protein (2.3.1).

In *E. coli* two classes of signal peptidases have been identified, signal peptidases I and II, both of which are essential for viability. Signal peptidase I, or Lep, is the product of the *lep* (leader peptidase) gene and is involved in the processing of the majority of *E. coli* precursors. Signal peptidase II, the product

of the *lsp* gene, is responsible for the processing of glyceride modified lipoproteins.

2.4.5a SipS - a B. subtilis *type I signal peptidase*

Initial attempts to clone a *B. subtilis* signal peptidase using techniques dependent upon homologies between corresponding genes and proteins from *E. coli* were unsuccessful. However, a *B. subtilis* type I signal peptidase gene, *sipS*, was eventually cloned in *E. coli* by van Dijl *et al* (1992) who studied the ability of random *B. subtilis* chromosomal DNA fragments to promote processing of hybrid β-lactamase precursors *in vivo* (van Dijl *et al*., 1992). The *B. subtilis sipS* gene encodes a protein of 184 amino acids with a molecular mass of 21 kDa. SipS is much smaller than Lep from the Gram-negative organisms *E. coli* and *S. typhimurium* which are 323 and 324 amino acids long, respectively. Similar methods were employed to clone the type I signal peptidase genes of other *Bacillus* species. The genes from *B. amyloliquefaciens*, *B. licheniformis* and *B. caldolyticus* termed, *sipA*, *sipL* and *sipC*, respectively, have now been cloned by van Dijl *et al* (1995). The products of these genes exhibit 70% to 80% amino acid identity with SipS.

The overall homology between SipS and Lep is low which accounts for the inability to detect *sipS* using the *E. coli lep* gene as a probe. As a consequence of the smaller size of SipS, four regions found in the Lep protein of *E. coli* are absent from SipS and the similarities of Lep to SipS are confined to three non-contiguous regions of Lep (van Dijl *et al*., 1992). These regions of Lep, termed I, II and III, exhibit 27.5%, 34.7% and 39.1% identity with corresponding regions of SipS, respectively.

The similarities between specific regions of the proteins are not restricted solely to bacterial species as a certain degree of homology has been found between SipS and eukaryotic proteins involved in the processing of proteins at membranes (van Dijl *et al*., 1992). Examples of eukaryotic proteins with homology to SipS and Lep include the Sec11 subunit of the ER signal peptidase complex of *Saccharomyces cerevisiae*, the mitochondrial inner membrane protease 1 (Imp1) of *S. cerevisiae* and two subunits of the canine ER signal peptidase complex (Spc18 and Spc21). Further to the above, a comparison between SipS and Imp1 revealed a higher degree of homology between these two proteins when compared with the other bacterial signal peptidases. The homology was distributed throughout the entire length of the proteins and revealed 40% similarity and 26.6% amino acid identity (van Dijl *et al*., 1992).

The proposed model for the membrane topology of the *E. coli* Lep protein indicates that it has two transmembrane domains (H1 and H2) and a large periplasmic domain (H3), with both the N- and C-termini located on the periplasmic side of the cytoplasmic membrane (Fig. 2.4) (Dalbey and Wickner, 1987). Region I of SipS overlaps with the membrane spanning region of Lep H2, the region putatively involved in the catalytic activity of Lep (Bilgin *et al.*, 1990). SipS regions I and II correspond to regions of Lep which are exposed to the periplasm. On the basis of the similarities between SipS and Lep and the model by Dalbey, van Dijl *et al* proposed a model for the membrane topology of SipS (van Dijl *et al.*, 1992).

The model indicates that the N- and C-termini of SipS are located on the inside and outside of the cytoplasmic membrane, respectively (Fig. 2.4). Further, SipS has one transmembrane domain corresponding to region H2 of Lep, with the majority of the SipS protein located on the external side of the cytoplasmic membrane (van Dijl *et al.*, 1992). Removal of the cytoplasmic domain and transmembrane region H1 from Lep does not alter the catalytic activity (Dalbey and Wickner, 1987). Therefore, it seems that these regions are not essential for Lep function and may explain why corresponding regions are missing from SipS.

A similar situation exists with another component of the export machineries of *B. subtilis* and *E. coli*, namely SecE. In *B. subtilis* SipS, like SecE, has only one transmembrane domain (2.4.2c). The corresponding proteins from *E. coli*, Lep and SecE, have two and three transmembrane domains, respectively. In both cases the *B. subtilis* homologues are smaller in size and have only a single transmembrane domain. This observation may reflect evolutionary events during which non-essential domains of the proteins have been lost from *B. subtilis*, or it may simply indicate subtle differences between the export pathways of Gram-positive and Gram-negative bacteria.

Studies using site-directed mutagenesis of amino acid residues conserved throughout signal peptidases identified several amino acids which are important for the function of SipS (van Dijl *et al.*, 1995). In particular, two amino acid residues were implicated as conformational determinants of the SipS protein, namely aspartate-146 and arginine-84, and three residues were putatively identified as making up the active site of the enzyme on the outer surface of the cytoplasmic membrane; serine-43, lysine-83 and aspartate-153.

2.4.5b SipS, a non-essential but rate limiting factor of secretion ?

When SipS is overproduced in *B. subtilis* a concomitant increase in the rate of processing of secretory precursors of Gram-negative and Gram-positive

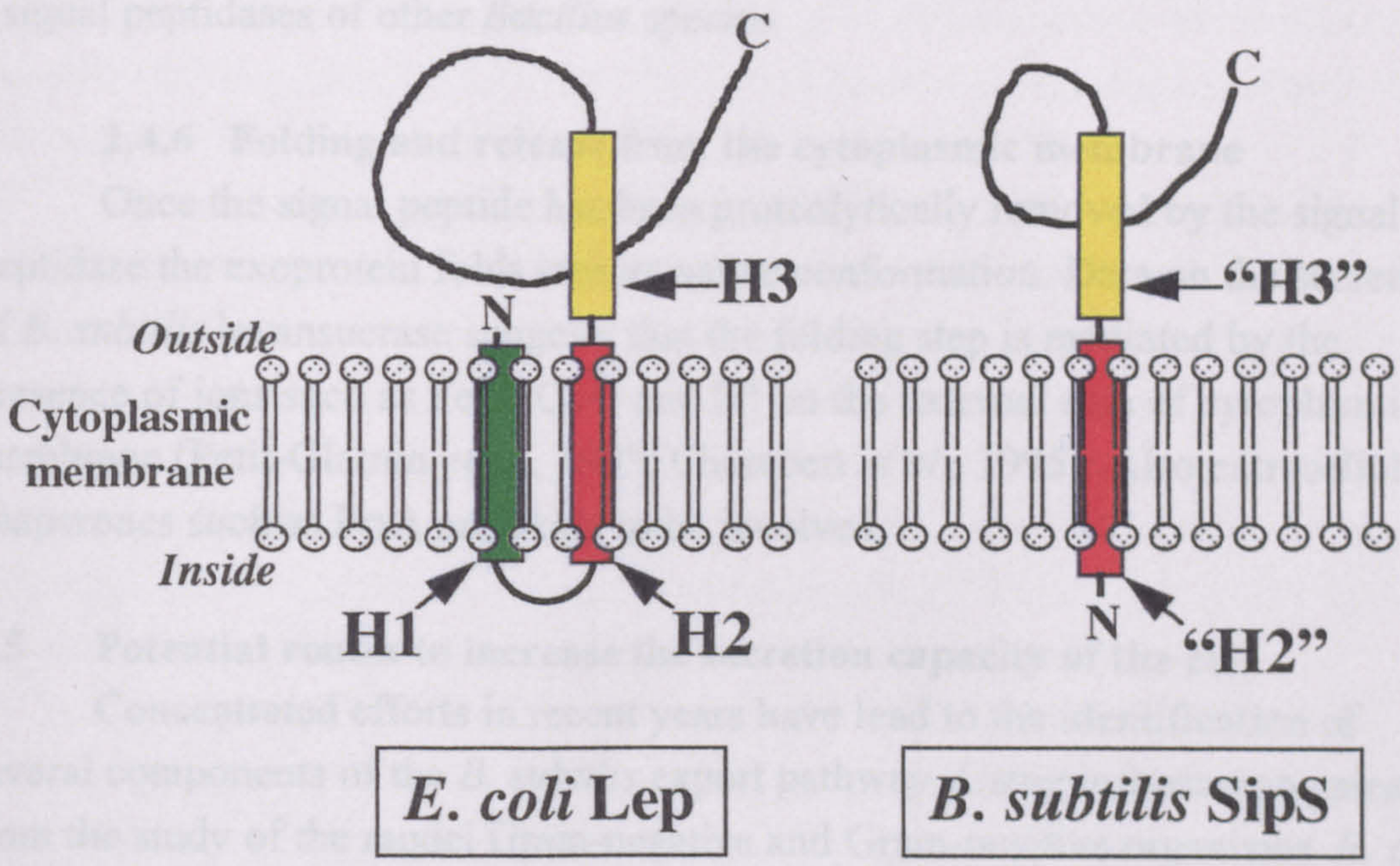


Fig. 2.4. The proposed membrane topology of type I signal peptidases of *E. coli* and *B. subtilis*.

origin is observed (van Dijl *et al.*, 1992). These findings suggest that SipS is rate limiting for protein export and increasing the number of copies of the *sipS* gene provides an additional way to increase the secretion capacity of the cell. *sipS* mutants remain viable and retain the ability to process precursors, albeit at a much reduced level (van Dijl *et al.*, 1992). Therefore, *B. subtilis* must possess at least one additional type I signal peptidase. Recently, a second type I signal peptidase from *B. subtilis* has been identified (Tjalsma *et al.*, 1995). This second type I signal peptidase, SipT, is able to process hybrid β-lactamase precursors in a similar way to that of SipS and has a high degree of similarity to SipS and type I signal peptidases of other *Bacillus* species.

2.4.6 Folding and release from the cytoplasmic membrane

Once the signal peptide has been proteolytically removed by the signal peptidase the exoprotein folds into its native conformation. Data on the secretion of *B. subtilis* levansucrase suggests that the folding step is mediated by the presence of ions such as Fe^{3+}, Ca^{2+} and H^{+} on the external side of cytoplasmic membrane (Petit-Glatron *et al.*, 1993; Chambert *et al.*, 1995). Also extracellular chaperones such as PrsA are likely to be involved.

2.5 Potential routes to increase the secretion capacity of the cell

Concentrated efforts in recent years have lead to the identification of several components of the *B. subtilis* export pathway. Using information gained from the study of the model Gram-negative and Gram-positive organisms, *E. coli* and *B. subtilis*, the processes by which proteins become extracellular are now relatively well understood. A detailed knowledge of the secretion components and the stages involved in the export process has important implications for the industrial exploitation of *B. subtilis* and other members of the genus.

Industrial production of proteins almost exclusively involves the high-level expression of the gene of interest. With such large quantities of protein entering the export pathway a situation could easily develop where it becomes saturated. Theoretically, secretion bottlenecks could develop at any stage in the export process, ranging from expression of exoprotein genes to the folding of the protein following translocation. Jamming of the secretion pathway is a major limitation in the use of bacteria for the export of proteins of industrial importance. With the increasing platform of knowledge relating to the secretion process, it may eventually prove possible to avoid secretion bottlenecks completely by increasing the copy numbers of the some or all of the genes encoding components of the export machinery. This would ultimately lead to the

development of a "cell factory" optimized for the high-level production of proteins of interest. The feasibility of this approach has already been demonstrated with overexpression of PrsA and SipS as discussed above.

2.6 The *B. subtilis* cell wall - introduction

The cell wall of *B. subtilis* lies on the external side of, and in close proximity to, the cytoplasmic membrane. The wall is a flexible and dynamically changing macromolecule which determines the size and shape of the cell. In *B. subtilis* and other Gram-positive organisms, the wall provides a physical barrier between the cytoplasmic membrane and the extracellular milieu, although certain members of the genus, such as *Bacillus brevis*, have a proteinaceous s-layer on the outer surface of the wall.

The cell wall is the structural component which serves to protect the protoplast from potentially dangerous swelling caused by the movement of water molecules by osmosis. In the absence of a wall the fragile cytoplasmic membrane would be susceptible to disruption due to the unrestricted inflow of water, leading ultimately to osmotic lysis. Because the wall of *B. subtilis* is directly exposed to the environment it is inherently involved in the interactions of the bacterium with its surroundings. As a consequence, all molecules or ions entering or leaving the cell must traverse and/or interact with the wall.

The *B. subtilis* cell wall is composed of a thick layer of peptidoglycan, a heteropolymer found only in prokaryotes (Rogers *et al.*, 1980), with covalently attached anionic polymers. The polymeric fabric of the wall forms a sacculus built up of many interconnected layers wrapped around the length and width of the cell. The structural organization, assembly and turnover of the cell wall of *B. subtilis* and other Gram-positive organisms has been reviewed recently (Archibald *et al.*, 1993).

The covalently attached anionic polymers, which can account for up to 50% of the weight of the wall, represent a high concentration of immobilized negative charge on the outside of the cell. The anionic polymers, teichoic acid and teichuronic acid, give the wall the chemical properties of an ion exchange matrix within which charged entities, such as proteins or ions, may exist in a state of equilibrium with constant association and dissociation. Molecules or ions with a high positive charge can be expected to interact more strongly with the wall and bind to it with higher affinity. In addition to the chemical properties of the wall, the layers of interconnected peptidoglycan chains give the wall a mesh-like structure with limited porosity. This alone may provide a physical barrier to the passage of larger molecules.

The *B. subtilis* cell wall not only gives the delicate protoplast a shield from the external environment but it also provides the cell with an external polyelectrolyte gel matrix which is required for many essential processes. For example, the combined physical and chemical properties of the wall maintain the necessary conditions for the activity of various membrane- and cell wall-associated enzymes. Further, the wall plays an important part in cellular growth, cell division and segregation of the chromosome.

2.7 Peptidoglycan, a heteropolymer exclusive to prokaryotes

To maintain the morphology of the bacterium and prevent osmotic lysis, the cell wall must be able to withstand the turgor pressure imposed upon it by the expanding cytoplasmic membrane caused by the influx of water by osmosis. The component of the wall responsible for its mechanical strength is peptidoglycan. The heteropolymeric peptidoglycan molecule is built up of linear chains of glycan which are connected together by short peptides. Every glycan chain is cross-linked by several peptides, each of which is in turn capable of cross-linking to a different glycan chain. This arrangement of rigid glycan chains cross-linked by short peptides produces a large macromolecule which is not only strong but has considerable elasticity. As a consequence, the wall can shrink or swell with changes in osmolarity, ionic strength or pH, without adversely affecting the underlying protoplast.

The glycan chains are assembled from repeating units of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc), which are covalently joined by β-(1-4) glycosidic linkages. In *B. subtilis* 168 the average biosynthetic glycan chain length is 100 disaccharide units (Ward, 1973). The cross-linking peptides are joined to the glycan chains by an amide linkage between the N-terminal L-alanine of the peptide and the carboxyl groups of the muramic acid residues (Fig. 2.5).

The N-terminal L-alanine residue of the cross-linking peptide is linked through its carboxyl group to the amino moiety of D-glutamic acid. The glutamic acid residue is then linked, *via* the γ carboxyl, to the amino group at the L-centre of diaminopimelic acid. Finally, the carboxyl group at the L-centre of diaminopimelic acid is linked to D-alanyl-D-alanine. During the cross-linking process transpeptidation reactions remove the D-alanine residue at the C-terminus of the "donor" pentapeptide and link the D-alanine at position 4 to the amino group of diaminopimelic acid on a different "acceptor" peptide chain. D-alanine carboxypeptidases may remove terminal D-alanine residues which destroys the

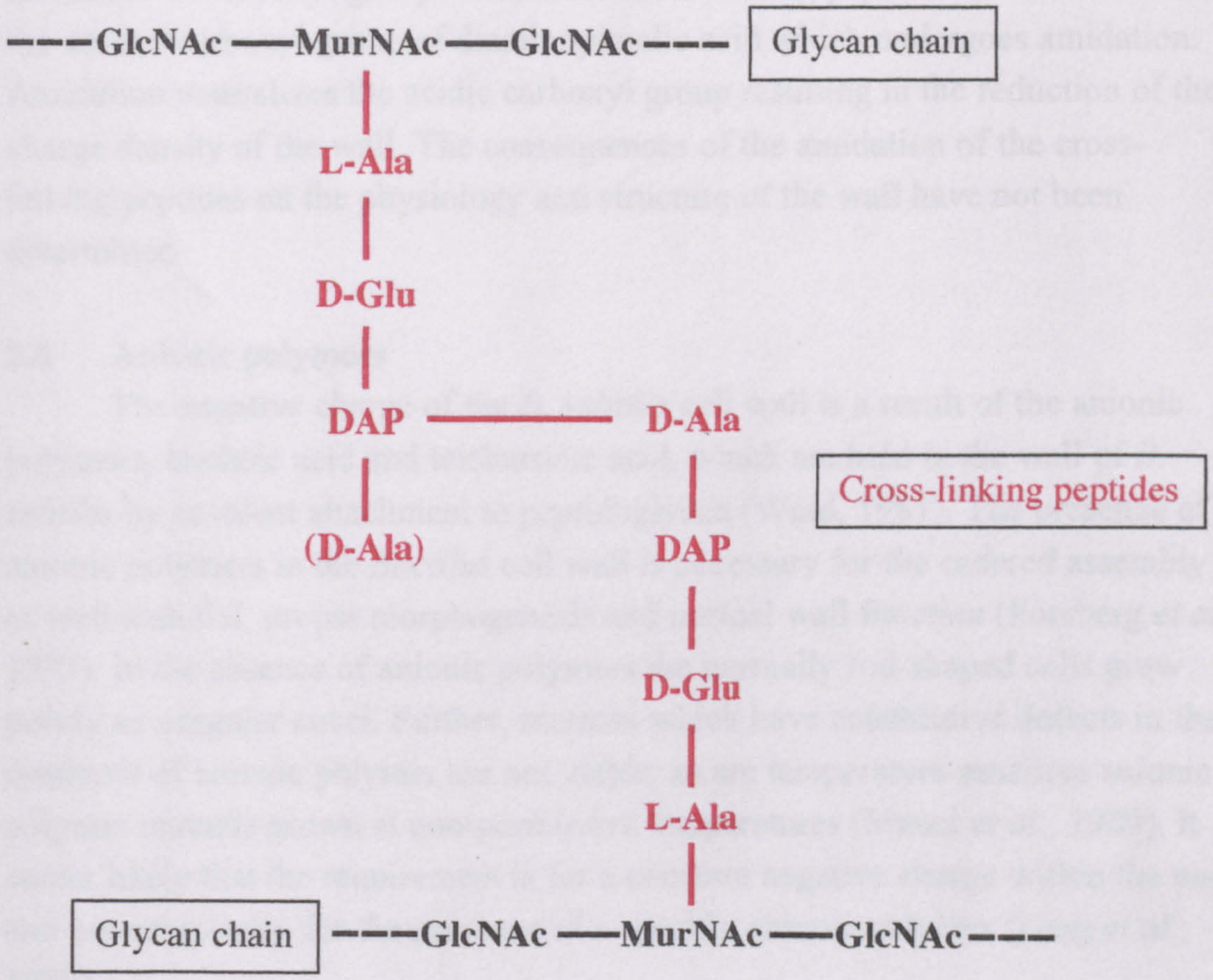


Fig. 2.5. The cross-linked structure of the peptidoglycan of *B. subtilis*.

ability of the peptide chain to acts as a donor. The action of carboxypeptidases may allow the cell to modulate the degree of cross-linking in the wall.

The charge density of peptidoglycan may be affected by the extent of the amidation of carboxyl groups within the cross-linking peptides. In *B. subtilis* it is the second carboxyl group of diaminopimelic acid which undergoes amidation. Amidation neutralizes the acidic carboxyl group resulting in the reduction of the charge density of the wall. The consequences of the amidation of the cross-linking peptides on the physiology and structure of the wall have not been determined.

2.8 Anionic polymers

The negative charge of the *B. subtilis* cell wall is a result of the anionic polymers, teichoic acid and teichuronic acid, which are held in the wall of *B. subtilis* by covalent attachment to peptidoglycan (Ward, 1981). The presence of anionic polymers in the *Bacillus* cell wall is necessary for the ordered assembly of wall material, proper morphogenesis and normal wall function (Forsberg *et al.*, 1973). In the absence of anionic polymers the normally rod-shaped cells grow poorly as irregular cocci. Further, mutants which have constitutive defects in the synthesis of anionic polymer are not viable, as are temperature sensitive anionic polymer mutants grown at non-permissive temperatures (Mauel *et al.*, 1989). It seems likely that the requirement is for a constant negative charge within the wall and not necessarily for the presence of a specific anionic polymer (Lang *et al.*, 1982).

The exact roles of the anionic polymers in the wall are unclear. The negatively charged polymers are known to interact with, and bind to, divalent metal ions such as Mg^{2+} and Ca^{2+} (Hepstinall *et al.*, 1970; Hughes *et al.*, 1973). This may have the effect of localizing a buffer of essential divalent cations close to the membrane which can be utilized by the cell as required. In a similar manner the metal binding capacity of the wall anionic polymers may act as a sieve to stop potentially harmful toxic metal ions coming into to direct contact with the cytoplasmic membrane. In addition, the teichoic acids of *B. subtilis* interact specifically with N-acetylmuramoyl-L-alanine amidase, the main autolysin of this strain (Herbold and Glaser, 1975b) (2.12.1). This interaction may be responsible for the localization of this enzyme in the wall and/or in modulating its activity during growth and wall turnover.

Teichoic and teichuronic acids contain negatively charged groups that dominate the overall polyelectrolyte gel of the wall and therefore, contribute to surface charge and specific surface properties. Teichoic acids contain charged

phosphodiester groups in the repeating units of the polymer while the charged carboxyl groups of the uronic acid residues of teichuronic acids are responsible for the negative charge (Fig. 2.6).

The wall is made up of a peptidoglycan-anionic polymer complex with the anionic polymers attached throughout the thickness of the wall. This complex is formed at the inner surface of the wall cylinder and moves to the outer surface as a consequence of growth and wall turnover. The anionic polymer composition of the wall changes with growth conditions (2.10) but the overall negative charge of the wall remains relatively constant (Lang *et al.*, 1982). However, even with the maintenance of a constant negative charge within the wall, differences in the polymer content can result in subtle changes in the wall and the surface chemistry of the cell.

The anionic polymers are covalently linked to the N-acetylmuramic acid of the glycan chain by the hydroxyl group at position C-6 of the muramyl residue. Teichoic acids are attached to the muramyl residue by a specific "linkage unit" however, no such linkage unit has been identified in the attachment of teichuronic acid to the glycan chain. The type of anionic polymer present in the wall of *B. subtilis* is dependent upon the strain and the conditions under which the bacteria are grown (2.10).

2.8.1 Teichoic acid

B. subtilis Marburg (NCTC 3610) and *B. subtilis* 168 have walls with a 1,3-linked glycerol teichoic acid (Fig. 2.6) when grown in phosphate replete medium (Glaser and Burger, 1964; Birdsell *et al.*, 1975). In contrast, the teichoic acid of *B. subtilis* W23 is a ribitol phosphate polymer.

Glycerol teichoic acids may be substituted by α-glucosyl or D-alanyl esters at the hydroxyl group at the C-2 position of the glycerol residue. The alanyl ester linkage is extremely labile at pH values above 7.0 due to the adjacent phosphate and hydroxyl groups (Shabarova *et al.*, 1962). This lability at alkaline pH determines the degree to which the glycerol teichoic acid is substituted by alanine. Under physiological conditions, the amino group of the alanine is protonated. This has the effect of neutralizing adjacent phosphate groups and consequently decreasing the negative charge of the cell wall. The ribitol teichoic acid of *B. subtilis* W23 is also substituted by a D-alanyl ester at position C-2.

Teichoic acids are covalently attached to the muramic acid residue of the glycan chain of *B. subtilis* by a "linkage unit" (Coley *et al.*, 1978). The linkage unit of *B. subtilis* W23 is composed of two glycerolphosphate residues attached to N-acetylmannosaminyl-N-acetylglucosaminyl-phosphate. The disaccharide is

linked to muramic acid by a sugar-1-phosphate linkage. The way in which the teichoic acid of *B. subtilis* 168 is linked to peptidoglycan has not been determined but it is assumed to attach in a similar way to that of W23. In addition, analysis of the linkage of teichoic acids to the peptidoglycan of various *Bacillus* species suggests that the linkage unit found in *B. subtilis* W23 is relatively well conserved.

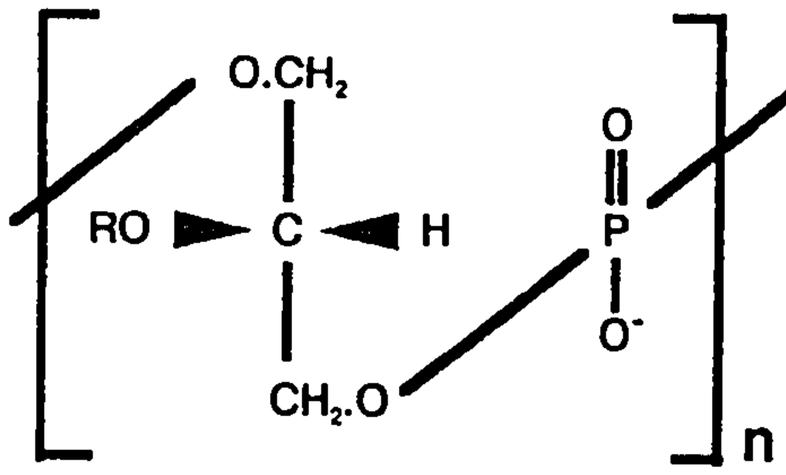
2.8.2 Teichuronic acid

When most strains of *B. subtilis* are grown under conditions of phosphate limitation, the anionic polymer in the wall consists mainly of teichuronic acid (2.10). No data is available about the structure of the teichuronic acid from *B. subtilis* 168, however, data does exist for the teichuronic acid from *B. subtilis* W23 which consists of linear chains of repeating disaccharide units (Fig. 2.6) (Wright and Heckels, 1975). Although the teichuronic acid molecule is covalently attached to the peptidoglycan of *B. subtilis* W23, no linkage unit similar to that for teichoic acids has been identified. The teichuronic acid of *B. subtilis* AHU 1031 is made up of a chain of 12 to 15 repeating units which terminates with an N-acetylglucosamine group which is likely to be linked to the muramic acid residue of peptidoglycan by a sugar-1-phosphate linkage (Yoneyama *et al.*, 1984). It is possible that the teichuronic acid of strain 168 is linked to peptidoglycan in a similar manner.

2.8.3 Lipoteichoic acid

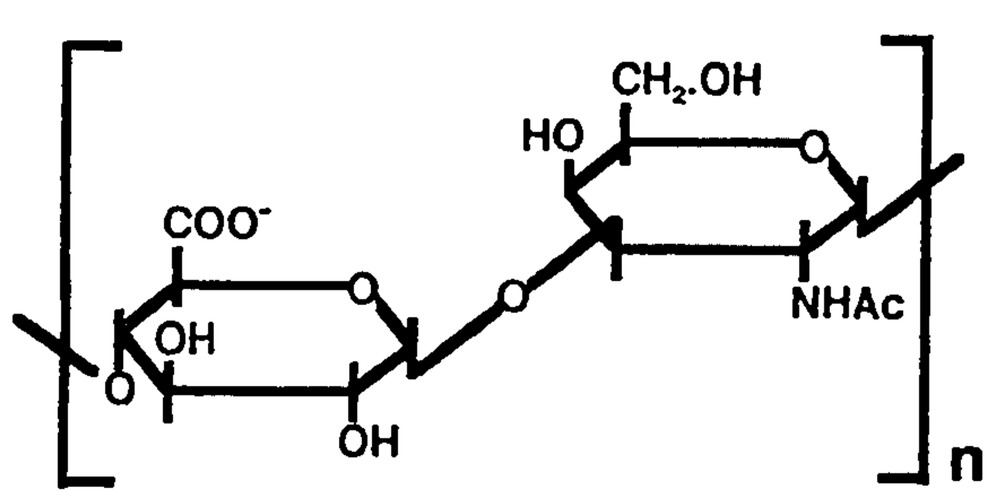
In addition to teichoic and teichuronic acids, several Gram-positive organisms have another anionic polymer on the external side of the cytoplasmic membrane, namely lipoteichoic acid. The lipoteichoic acid of *B. subtilis* is composed of a polyglycerol chain which is made up of 24 to 33 phosphodiester-linked units. The chain is attached by the terminal phosphate to 3-gentiobiosyl diglyceride which is responsible for cytoplasmic membrane anchoring (Iwasaki *et al.*, 1986). The polyglycerol-phosphate chains of lipoteichoic acid and wall teichoic acid differ in their stereochemical properties, reflecting the different biosynthetic origins of the polymers (Koch and Fischer, 1978). The biological significance of the opposite stereochemistry of lipoteichoic and teichoic acids has not been determined.

The lipid modification of lipoteichoic acid ensures that it is mainly associated with the cytoplasmic membrane but it may also be present in the wall or excreted into the culture supernatant. The lipoteichoic acid present in the wall may not be a specific wall-associated form but may simply be in the process of



The glycerol teichoic acid of *B. subtilis* 168

R = glucosyl residue



The teichuronic acid of *B. subtilis* W23

Fig. 2.6. The repeating units of the anionic polymers of *B. subtilis*.

being excreted to the external environment. Lipoteichoic acid can be released from the membrane by the addition of endogenous deacylating lipase (Kessler *et al.*, 1979) and therefore, it is likely that the acyl groups of the lipid moiety inserts into the membrane and intercalates with hydrophobic groups therein. Further, the addition of metal chelators such as ethylenediaminetetraacetic acid (EDTA) causes lipoteichoic acid to be released (Hughes *et al.*, 1973) suggesting that the insertion of lipoteichoic acid into the cytoplasmic membrane is stabilized by the presence of divalent cations.

2.8.4 The location of teichoic and teichuronic acid in the cell wall

Teichoic and teichuronic acids are attached to peptidoglycan throughout the wall and are not localized within particular regions of the wall cylinder. The anionic polymers are joined to the peptidoglycan chains at the inner surface of the wall at the site at which the peptidoglycan chains are assembled. The anionic polymer-peptidoglycan complex then moves from the inner to the outer surface of the wall cylinder as a consequence of cell growth and subsequent wall turnover.

Ligand binding studies demonstrated that a large amount of the teichoic acid in the walls of *B. subtilis* 168 is exposed at the surface and can participate in ligand binding (Birdsell *et al.*, 1975; Doyle *et al.*, 1975). Further, investigation of the binding of teichoic acid-specific bacteriophages and concanavalin A to walls of *B. subtilis* revealed that approximately half the anionic polymer content of the wall is located below the surface (Archibald and Coapes, 1976; Archibald, 1976; Anderson *et al.*, 1978). In these studies *B. subtilis* cells were grown under phosphate limitation resulting in a teichuronic acid-containing wall. When the phosphate-limited cells were switched to a medium with excess phosphate the newly incorporated teichoic acid was inaccessible to bacteriophage and concanavalin A until it moved to the outer surface of the wall cylinder as a consequence of growth.

2.9 Cell wall assembly and turnover

Before assembly, the component polymers of the *B. subtilis* cell wall are synthesized in different locations within the cell. In the first stage the nucleotide precursors are synthesized in the cytoplasm by enzymes which may be loosely associated with the cytoplasmic membrane (Hancock and Baddiley, 1985). The units are then transferred from the nucleotides to a lipid carrier in the membrane, undecaprenol phosphate. Following polymerization of peptidoglycan and anionic polymer by membrane-bound enzyme complexes, the units are incorporated into

the existing wall material by transpeptidation reactions at the cell wall-cytoplasmic membrane interface. In growing bacteria, newly synthesized anionic polymer molecules become attached only to newly synthesized peptidoglycan (Mauck and Glaser, 1972) which is undergoing cross-linking into the existing wall cylinder.

Cylindrical extension of the wall occurs by the incorporation of new wall material at many sites to cover the entire inner surface of the cylindrical wall (Archibald, 1976). This new material moves through the wall cylinder in an inside-to-outside manner and is replaced by newer wall material at the inner surface. As the wall material proceeds towards the surface it is progressively "stretched" until, when at the surface, it is cleaved, or turned over, by autolytic enzymes which ensure that the thickness of the wall cylinder remains relatively constant (Fan *et al.*, 1975; Archibald and Coapes, 1976; Koch, 1983; Merad *et al.*, 1989)

Following cylindrical extension, septation proceeds from an annular zone which results in a flat disk of wall material bisecting the cell. Autolysins act in a controlled manner to cleave the septal disk symmetrically, forming two daughter cells. The mode of autolytic regulation at this site is not well defined. After separation the septal disks form one of the poles of each daughter cell. During growth new wall material is incorporated into the poles and like the cylindrical walls, this occurs by an inside-to-outside mechanism. However, the incorporation of wall material at the poles occurs at a reduced rate when compared to the cylindrical wall and therefore, the polar caps turn over less rapidly (Archibald and Coapes, 1976; Mobley *et al.*, 1984; Clarke-Sturman *et al.*, 1989).

2.10 The effect of growth conditions on wall composition and structure

As the cell grows the peptidoglycan-anionic polymer complex is constantly replaced by new material which is assembled at the inside of the wall cylinder by membrane associated lipid-linked enzymes. New wall material moves through the wall towards the external surface and this inside-to-outside growth reflects the dynamically changing nature of the cell wall. Older material at the outside of the wall cylinder is constantly degraded in a controlled manner by the action of cell wall-associated autolysins.

The best studied example of the way in which growth conditions can alter the composition of the cell wall is the change that occurs in the wall when *B. subtilis* is grown under phosphate limitation. During growth in phosphate replete medium the major anionic polymer in the cell wall is the phosphate-containing teichoic acid. However, when the level of phosphate in the growth medium

becomes limiting the cells are able to switch the anionic polymer from teichoic acid to the non-phosphate containing teichuronic acid (Ellwood and Tempest, 1969; Ellwood and Tempest, 1972). At any point during phosphate limitation the total amount of anionic polymer within the wall, *i.e.* teichoic plus teichuronic acid, remains constant (Lang *et al.*, 1982).

During phosphate limitation the carboxyl groups of teichuronic acid maintain a relatively constant negative charge within the wall in the absence of teichoic acid. This switch allows the cell to use limited supply of phosphate for more essential phosphate-containing biological molecules, such as nucleic acids. The amount of teichoic acid incorporated into the walls in *B. subtilis* cells grown under phosphate limitation is regulated by modulating the activity of the enzymes required for the synthesis of the lipid-linked linkage unit, and not the main chain of the teichoic acid molecule (Cheah *et al.*, 1982; Hancock, 1983; Hancock, 1985).

In addition to the anionic polymer content of the wall, the degree of alanylation of teichoic acid can also be affected by the conditions under which *B. subtilis* and *B. licheniformis* are grown (Ellwood and Tempest, 1972). The effect of alanine esterification in the teichoic acids of *B. subtilis* is often underestimated or sometimes even overlooked. The amino group of the alanine is capable of neutralizing the adjacent phosphate group of teichoic acid and therefore, reduces the negativity of the polyelectrolyte gel of the wall. As a consequence of this, the capacity of the wall to bind cations is greatly reduced (Hepstinall *et al.*, 1970). Additionally, a decrease in the negative charge of the wall will reduce the electrostatic interactions between the teichoic acid molecules and therefore, reduce the separation between different teichoic acid chains caused by electrostatic repulsion.

Like wall teichoic acid, lipoteichoic acid can be substituted by D-alanyl esters which are labile in alkaline conditions. The degree of alanylation of lipoteichoic acid affects the binding capacity and porosity of the wall in a similar manner to that of wall teichoic acid. Further, the alanine content of lipoteichoic acid determines the degree to which this anionic polymer can interact with, and modulate the activity of, autolysins present in the cell wall (Fischer *et al.*, 1981a; Fischer *et al.*, 1981b). Highly alanylated lipoteichoic acid is unable to inhibit the activity of autolytic enzymes which may have drastic consequences for the bacterial cell (2.12.2).

To summarise, the conditions under which *B. subtilis* and other members of the genus are grown have a marked effect on the composition of the cell wall.

Changes in the composition of the wall can, in extreme situations, have a dramatic affect on wall structure and the specific surface properties of the cell.

2.11 Wall-associated proteins of *B. subtilis*

In addition to anionic polymers, *B. subtilis* and other members of the genus have several specific proteins associated with the cell wall. In *B. subtilis* 168 at least twelve major vegetative wall-associated proteins (WAP's) have been identified which can be extracted from cell walls using high salt or denaturing conditions (Studer and Karamata, 1988). One group of major wall-associated proteins are the autolysins which are involved in cell wall turnover. Of the twelve identified WAP's, four are autolysins and a fifth is the amidase modifier protein (2.12) (Foster, 1992; Foster, 1993). The functions of the remaining WAP's are as yet undetermined.

The structural gene for the 109 kDa WAP from *B. subtilis* was cloned and sequenced in an attempt to investigate the properties of this protein (Foster, 1993). Western blotting, using antisera raised against this protein, revealed the presence of two other cross-reacting proteins in wall and supernatant fractions with molecular masses of 220 kDa and 58 kDa. It was subsequently determined that the three related WAP's were processed forms of a 258 kDa precursor, termed WapA, which is the largest protein known to be naturally secreted by *B. subtilis*. It is not known whether processing of WapA occurs during or after translocation but the processed forms can be associated with the wall or free in the supernatant.

Peptidoglycan binding domains of proteins from Gram-positive organisms are characterized by the presence of directly repeated sequences which do not necessarily exhibit inter-protein sequence conservation (Garcia *et al.*, 1990; Kuroda *et al.*, 1992; Lazarevic *et al.*, 1992). The N-terminal region of WapA is common to all forms of the protein and contains three direct repeats of a 102 amino acid sequence. The three direct repeat sequences share 40% amino acid identity and insertional inactivation studies indicate that these repeats are responsible for wall binding. The N-acetylmuramoyl-L-alanine amidase of *B. subtilis* also has similar wall binding motifs located in its N-terminus. However, the exact structural determinants which mediate wall binding are not known.

Some of the WAP's of *B. subtilis* may also be found in the supernatant and it is likely that these "free" forms are released as a consequence of wall turnover. Apart from the proteins involved with cellular autolysis, the role of the WAP's in protein export, wall metabolism or other cell surface-associated phenomena remains undetermined. Studies on WapA from *B. subtilis* have not provided any

clues to their functions since mutants lacking this protein display no change in growth or physiology.

2.12 Autolysins are important cell wall-bound exoenzymes

During growth new cell wall material is incorporated at the inner surface of the wall cylinder and simultaneously degraded at the outer surface. Wall turnover at the outer surface is controlled by the activity of cell wall-associated enzymes termed autolysins. Autolysins function to ensure that the thickness of the wall cylinder remains relatively constant. In addition to wall turnover, autolysins are responsible for the cleavage of the septum during cell division and are involved in several post-exponential phase processes such as sporulation, germination and the development of genetic competence.

The major autolysin of *B. subtilis* is N-acetylmuramoyl-L-alanine amidase, which hydrolyzes the amide linkage between muramic acid of peptidoglycan, and the N-terminal alanine of the cross-linking peptide. The gene encoding this amidase from *B. subtilis*, *lytC* (*cwlB*), has been identified, cloned and sequenced (Kuroda and Sekiguchi, 1991; Margot and Karamata, 1992). The product of the *lytC* gene is 50 kDa in size and is functionally active as a dimer.

A second protein, the product of the *lytB* (*cwbA*) gene, associates with the amidase in the cell wall and modulates its activity by up to 3-fold (Herbold and Glaser, 1975b). This 80 kDa modifier protein alters the pattern of peptidoglycan cleavage by the amidase from a random to a more sequential process. In addition to the amidase, *B. subtilis* 168 has another autolysin present in the wall in significant quantities. This enzyme, endo-β-N-acetyl glucosaminidase, exists as a monomer with a molecular mass of 90 kDa and is the product of the *lytD* gene (Herbold and Glaser, 1975a; Margot *et al.*, 1991).

2.12.1 Autolysins bind to anionic polymers in the cell wall

The N-acetylmuramoyl-L-alanine amidase of *B. subtilis* can only bind to cell walls containing teichoic acid (Herbold and Glaser, 1975a). Teichoic acid-containing walls are required for the functional interaction of the amidase with its modifier protein and mutants deficient in this anionic polymer lack autolytic enzymes in their cell walls (Boylan *et al.*, 1972). It is likely that the failure to detect autolysins in the walls of these strains reflects the inability of the enzymes to bind to the wall in the absence of teichoic acid and not a defect in the synthesis, secretion or activity of the autolysins.

When teichuronic acid replaces teichoic acid, the presence of the modifier protein inhibits the activity of the amidase and seems to decrease the binding of

amidase to the cell wall (Herbold and Glaser, 1975a). Consequently, during phosphate limitation, the rate of wall turnover should decrease as teichoic acid is switched to teichuronic acid, however, this has yet to be demonstrated. The reason for the lack of binding of amidase to teichuronic acid is also unknown. Molecular models suggest that the spacing of negative charges in teichoic and teichuronic acid-containing walls is essentially the same and should not affect the binding of amidase or modifier protein. In contrast to *B. subtilis*, the autolysins of *B. licheniformis* will only bind to walls containing teichuronic acid and mutants lacking this polymer are resistant to autolytic activity and grow as long multiseptate filaments (Robson and Baddiley, 1977a; Robson and Baddiley, 1977b).

The ability of the amidase of *B. subtilis* to bind to cell walls is a function of the N-terminus, while the catalytic activity appears to be attributable to the C-terminus (Kuroda *et al.*, 1992; Lazarevic *et al.*, 1992). The N-terminus of the amidase contains three directly repeat sequences of about 55 amino acid residues which are thought to be involved in wall binding. The amidase modifier protein also has three similar motifs in its N-terminus. In addition to the directly repeated sequences, the amidase and its modifier are very basic, with isoelectric points (pI's) of 10.6 and 10.3, respectively (Lazarevic *et al.*, 1992). This is also likely to promote wall binding of these protein due to electrostatic interactions with anionic polymers (2.13.2).

2.12.2 The role and regulation of autolysins

Autolysins are responsible for cell wall turnover at the outer surface of the wall cylinder and also the separation of cells at the septum following cell division. As a consequence of the inside-to-outside growth of the wall, bonds within peptidoglycan become stretched as the peptidoglycan moves peripherally through the wall cylinder. It has been postulated that the autolysins are able to recognize and subsequently hydrolyze stretched peptide cross-links, thereby ensuring that it is only older wall which is degraded at a location distal to the cytoplasmic membrane (Koch *et al.*, 1985).

In *B. subtilis* the majority of wall turnover is achieved by the action of N-acetylmuramoyl-L-alanine amidase over the whole of the cell wall but at a reduced rate at the poles (Mauck *et al.*, 1971; Mauck and Glaser, 1972; Archibald, 1976; Archibald and Coapes, 1976; Sturman and Archibald, 1978; Mobley *et al.*, 1984). Approximately 50% of the cell wall is turned over each generation and in growing cultures of *B. subtilis* up to one third of the total wall

material present may exist as a soluble form liberated from the cell by the action of autolysins.

The cell wall curtails the expansion of the cytoplasmic membrane due to the unrestricted flow of water. Uncontrolled degradation of the wall will result in cell lysis and therefore, the activity of autolysins must be strictly controlled and restricted to the outermost layers of the wall only. The control of the activity of extracellular enzymes is a particular problem and *B. subtilis* regulates the action of autolysins in a number of ways.

An important autolytic regulatory mechanism seems to be associated with the energized cytoplasmic membrane since conditions that inhibit energy generation promote autolytic activity (Jolliffe *et al.*, 1981). *B. subtilis* has several other mechanisms which seem to restrict the hydrolysis of peptidoglycan to the outer layers of the wall such as the inhibitory affect of lipoteichoic on autolytic activity (Fischer *et al.*, 1981a; Fischer *et al.*, 1981b), the presence of a low pH close to the membrane and also the selective hydrolysis of stretched bonds by autolysins (Koch *et al.*, 1985). Additionally, the ionic environment (Cheung and Freese, 1985) and the activity of extracellular proteases also appear to be important in the control of autolytic activity in *B. subtilis* (Jolliffe *et al.*, 1980; Coxon *et al.*, 1991).

2.13 The potential influence of the cell wall on protein secretion

The outer membrane of Gram-negative bacteria limits the extent to which exoproteins can be secreted directly into the culture supernatant and the absence of this structure in *Bacillus* species provides one less barrier for certain secretory proteins to cross on their way to the external environment.

Following translocation, secretory proteins can remain associated with the cytoplasmic membrane, the cell wall or take up a truly extracellular location in the culture supernatant. The location of secreted proteins is dependent upon the nature of each individual exoprotein and examples from each category are PrsA, N-acetylmuramoyl-L-alanine amidase and extracellular alkaline protease (subtilisin), respectively. Proteins destined for the culture supernatant must traverse the cell wall as the last stage of the export process and therefore, interactions between the cell wall and secretory proteins must occur. The nature of such interactions is likely to be affected by the structural properties of the exoproteins and also those of the cell wall.

It has been suggested for *B. amyloliquefaciens* that proteins are translocated relatively rapidly across the cytoplasmic membrane into a pool and then move at a reduced rate across the wall (Gould *et al.*, 1975). In this case the

wall determines the rate at which proteins are released into the culture supernatant. In addition, Archibald (1985) reported a delay in secretion of pulse-labelled proteins from chemostat cultures of *B. subtilis*, suggesting that protein secretion may be coupled in some way to cell wall assembly and turnover. It is possible that proteins are picked up and transported through the wall as a consequence of growth and subsequently released at the outer surface as the wall is turned over. The *B. subtilis* cell wall is known to completely retard the secretion of some proteins, *e.g.* human serum albumin, to such a degree that this protein can only be secreted when the wall has been completely removed (Saunders *et al.*, 1987). Hence, it is clear that the wall is an important factor in protein secretion and one that must be negotiated by all exoproteins before they can take up residence in the supernatant.

By virtue of the immobilized negative charge, the cell wall has the effect of binding and concentrating cations. The wall concentrates Ca^{2+} ions approximately 100-fold, when compared to the culture supernatant, to maintain a high concentration of Ca^{2+} on the outside of the cytoplasmic membrane (Bevridge and Murray, 1976; Petit-Glatron *et al.*, 1993). The resultant high local Ca^{2+} concentration is sufficient to trigger the folding of *B. subtilis* levansucrase and it has been proposed that the cell wall acts as extragenic factor for secretion by maintaining a concentration of metal ions such as Ca^{2+} which promotes rapid folding following translocation (Petit-Glatron *et al.*, 1993).

2.13.1 The molecular sieving properties of the cell wall

The cell wall may not be freely permeable to all proteins due to the physical nature of the peptidoglycan-anionic polymer matrix. The cross-linked structure of peptidoglycan determines the porosity of the wall and the consequential molecular sieving properties can be expected to retard the passage of larger proteins more than smaller ones. Physical studies have indicated that the *Bacillus megaterium* wall may not be freely permeable to molecules with a radius of more than 1.1 nm (Scherrer and Gerhardt, 1971), but the significance of this estimation for protein secretion in *B. subtilis* is difficult to predict.

2.13.2 The cell wall as a cation-exchanger

The immobilized negative charge contributed by anionic polymers gives the cell wall the properties of a cation exchanger. Proteins entering and making their way through the wall cylinder are likely to interact with the anionic polymers. The strength of the resulting electrostatic interaction is dependent upon the net charge of the protein and also the negative charge of the wall. Little work

has been carried out on the affect of exoprotein charge on protein secretion but, on the basis of electrostatic attraction, exoproteins with an overall positive charge would be expected to associate strongly with the anionic polymers within the wall. This may have the affect of slowing their passage through the wall to such an extent that they move towards the surface as a consequence of growth and turnover only. Additionally, the ionic strength of the culture medium can also be expected to affect the binding as cations in solution will compete with basic residues for binding to the anionic polymers.

In contrast, negatively charged proteins may be forced across the wall as a consequence of electrostatic repulsion between the proteins and anionic polymers. This may increase the rate at which acidic proteins traverse the wall and are released into the culture supernatant. Conversely, following translocation, acidic proteins may be prevented from entering the wall cylinder completely due to the electrostatic repulsion.

When a protein is placed in a pH gradient it will migrate to a point at which the net charge of the protein is zero. The pH at which a protein has no net charge is termed the pI and the higher the pI value the more positively charged (*i.e.* basic) the protein is. Investigation of the pI values of pulse-labelled exoproteins produced by *B. subtilis* W23 revealed that all the rapidly exported exoproteins had pI values of between 6.5 and 7.0 (Coxon, 1990). At neutral pH these exoproteins would have little or no net charge and would therefore not be expected to interact with the cell wall to a significant extent. This may represent an evolutionary adaptation used by *Bacillus* species which tailors exoproteins in such a way to minimize interactions with the cell wall and promote unhindered egress from the cell envelope.

2.14 The affect of growth conditions on the passage of exoproteins through the cell wall

As described in previous sections, the *B. subtilis* cell wall is a dynamically changing macromolecule, the structure of which is affected in various ways by changes in growth conditions. The structure and composition of the wall, and the anionic polymers within it, can be affected by the pH, ionic strength or osmolarity of the culture medium.

The alanyl substituent of wall teichoic acid (and lipoteichoic acid) neutralizes the negative charge of the adjacent phosphate group. The degree of alanylation is affected by growth conditions which, in turn, has important implications for the passage of proteins through the cell wall due to the consequential affects on wall porosity and cation binding. If, for example, the

alanyl content of teichoic acid is reduced by an increase in the pH of the growth medium, the overall negative charge of the wall increases. This would cause corresponding increases in wall porosity and also the affinity of the wall for positively charged molecules or ions.

In a similar manner, the ionic strength of the growth medium can influence the porosity of the cell wall. Under conditions of low ionic strength, increased electrostatic repulsion between anionic polymers increases the separation between peptidoglycan strands with a consequential increase in wall porosity. Conversely, conditions of high ionic strength will cause the wall to adopt a more condensed structure and both of these effects are likely to influence the passage of proteins through the matrix of the cell wall. The ability of the wall to adjust in response to changes in the ionic strength of the medium reflects the elastic nature of the wall, which in turn is a function of the structure of the peptidoglycan-anionic polymer complex. Additionally, the teichoic/teichuronic switch, which occurs when the cell is phosphate limited, may also affect the degree to which proteins interact with the cell wall although this has not been investigated in any detail.

Members of the genus *Bacillus* secrete proteins efficiently and at a high level. Extracellular enzymes such as α-amylase and alkaline protease are exported in significant amounts and the cell wall does not appear to be a barrier to their passage. Other native *B. subtilis* exoproteins (*e.g.* WAP's) are bound tightly but non-covalently to the cell wall and in several of these proteins directly repeated wall binding motifs have been identified in the N-terminal domains. Whether these proteins have a function in wall assembly and/or turnover has not been determined. Attempts to use *B. subtilis* for the secretion of heterologous proteins such as human serum albumin has met with limited success due, at least in part, to the presence of the wall.

It is clear that the Gram-positive cell wall is an important determinant in protein secretion as its cross-linked structure and electrostatic properties have the potential to promote a high degree of interaction with secretory proteins. Consequently, the composition of the cell wall, which is influenced markedly by growth conditions, is likely to have a considerable affect on the extent of this interaction and, subsequently, the extent to which the passage of a protein is retarded by the wall.

2.15 The extracellular proteases of *B. subtilis*

When a culture of *B. subtilis* runs out of nutrients it stops growing exponentially and moves into the stationary phase. At this time a battery of

degradative enzymes are secreted into the culture supernatant. These enzymes perform a scavenging function in an attempt by the cell to salvage any remaining nutrients from the environment before it is driven to sporulation. Amongst the enzymes produced are extracellular proteases which break down polypeptides into oligopeptides and amino acids which can be used as carbon and nitrogen sources. The regulatory mechanism controlling the synthesis of extracellular proteases is complex and poorly understood, however, at least nine genes are known to be involved (Strauch *et al.*, 1989).

The presence of extracellular proteases in the culture supernatant is likely to prove problematic when using *B. subtilis* as the host for the production of heterologous proteins (2.16). This is because, whilst native *B. subtilis* exoproteins have evolved to be relatively resistant to the action of such enzymes, proteins from other origins are not. As a consequence, secreted heterologous proteins can be rapidly degraded, severely affecting the quality and quantity of recoverable product.

Two major classes of extracellular proteases are found in *B. subtilis*, serine proteases and metalloproteases. Serine proteases have a serine residue located within the active site and can be inhibited by hydroxyl-reactive organofluorides such as phenylmethylsulphonyl fluoride (PMSF). Metalloproteases require the presence of divalent metal cations for activity and are consequently inhibited by chelators of these ions such as EDTA. To date the gene for seven extracellular proteases from *B. subtilis* have been cloned and characterized.

The genes for the identified *B. subtilis* extracellular proteases are scattered throughout the chromosome and their products are described below. The two major proteases are the alkaline serine protease, AprE (subtilisin), and the neutral or metalloprotease, NprE. The properties of these proteases and the others identified to date, have been reviewed extensively (He *et al.*, 1991; Pero and Sloma, 1993). In addition to proteolytic activity, serine proteases exhibit a specific esterase activity that is not present in metalloproteases.

2.15.1 The *aprE* gene

The gene encoding the major extracellular alkaline serine protease or subtilisin, *aprE*, was the first protease gene to be cloned from *B. subtilis* (Stahl and Ferrari, 1984; Wong *et al.*, 1984). The *aprE* gene is translated to produce a preproenzyme comprising a signal peptide of 29 amino acids, a pro-region of 77 amino acids and the mature protein of 275 amino acid residues (Wong and Doi, 1986).

aprE expression is at its maximum during the early stages of sporulation and is regulated by a complex mechanism which involves several other genes including *spo0A*, *sacU*, *sacQ*, *sacV* and *abrB*. *aprE* is transcribed from a promoter recognized by the vegetative sigma factor, σ^A, and not from a post-exponentially active sigma factor (Park *et al.*, 1989).

Subtilisin is synthesized with a pro-region which is autocatalytically removed and the self-processing reaction is necessary for the release of the protein from the cytoplasmic membrane (Power *et al.*, 1986). In addition, the pro-region acts as an intermolecular chaperone and guides the folding of the mature protein into the active form (Ohta *et al.*, 1991). Exogenously added pro-sequence complements the folding of denatured subtilisin into the active form of the enzyme (Zhu *et al.*, 1989), confirming the role of this region in folding.

2.15.2 The *nprE* gene

The product of the *nprE* gene, the major neutral protease, is also synthesized once exponential growth has ceased. NprE is a metallo-endopeptidase which has a single zinc ion located at its active site. In addition to the zinc ion, as many as four calcium ions may be involved in stabilizing the protein. Cloning and sequencing of *nprE* have shown that its product is also produced as a prepropeptide (Yang *et al.*, 1984; Toma *et al.*, 1986). The protein consists of a signal sequence of 27 amino acids, an unusually long pro-region of 194 amino acids and a mature enzyme of 300 amino acid residues.

Like subtilisin, the expression of the *nprE* gene is under the control of numerous *spo0* genes and other genes involved in the regulation of extracellular proteins, such as *sacQ* and *sacU*. Many different genes are involved in the control of the expression of *nprE* (and *aprE*) but the control mechanism appears to be complex and the exact details remain unknown. The pro-region of NprE is required for the production of an active enzyme (Takagi and Imanaka, 1989) and it is assumed that the long propeptide is involved in the folding and autocatalytic release of mature neutral protease from the cytoplasmic membrane in a similar manner to that of subtilisin (Wandersman, 1990).

2.15.3 The *bpf* (*bpr*) gene

The product of the *bpf* gene, bacillopeptidase F, is an extracellular serine protease with relatively high esterolytic activity. The *bpf* gene has been cloned and sequenced and encodes a protein of 1433 amino acid residues (Wu *et al.*, 1990; Sloma *et al.*, 1990b). Bacillopeptidase F has a signal peptide of 30 amino acid residues and a putative propeptide of 164 residues. The mature protein is

made up of 1239 amino acid residues and has a calculated molecular mass of 135 kDa however, the protein is subjected to considerable C-terminal processing and forms of the enzyme above 50 kDa are not detectable in culture supernatants (Roitsch and Hageman, 1983; Wu *et al.*, 1990; Sloma *et al.*, 1990b).

Sequence comparisons of the amino acid sequence of bacillopeptidase F and subtilisin showed 30% similarity with a particularly high degree of similarity around residues known to be involved in the active site of subtilisin (Wu *et al.*, 1990; Sloma *et al.*, 1990b). In common with the other extracellular proteases, bacillopeptidase F is temporally expressed at, or shortly after, the onset of stationary phase.

2.15.4 The *mpr* gene

The construction of *B. subtilis* strains lacking major extracellular proteases allowed a number of minor proteases to be identified. A previously masked minor metalloprotease, Mpr, was isolated from a strain of *B. subtilis* lacking functional *aprE*, *nprE* and *bpf* genes (Rufo *et al.*, 1990). The *mpr* gene was cloned using an oligonucleotide probe designed on the basis of the N-terminal sequence of Mpr and was shown to code for a protein of 313 amino acid residues with a molecular mass of 24 kDa (Sloma *et al.*, 1990a). Mpr is synthesized as a preproenzyme which has a signal peptide of 34 amino acid residues and a propeptide of 58 residues. Additionally, four cysteine residues were identified in the mature protein which are capable of forming disulphide bonds, a property which is not found in the other *B. subtilis* extracellular proteases.

2.15.5 The *epr* gene

The product of the *B. subtilis epr* gene is a minor serine protease. Epr is synthesized as a preproenzyme with a signal peptide of 27 amino acid residues, a propeptide of 76 residues and a mature region of 542 residues (Sloma *et al.*, 1988; Bruckner *et al.*, 1990). Although the calculated molecular mass of the mature protein is 58 kDa, Epr is found in the culture supernatant in multiple forms with masses ranging from 34 kDa to 40 kDa. This suggests that, like bacillopeptidase F, Epr undergoes considerable processing. Deletion analysis of the C-terminal domain of Epr has indicated that this region is not essential for gene expression, secretion or enzymatic activity (Sloma *et al.*, 1988; Bruckner *et al.*, 1990).

A large number of the amino acid residues in the C-terminus are basic residues and confer an overall positive charge to this region of Epr. Further, there is a partially homologous region of 44 residues which is directly repeated five

times (Sloma *et al.*, 1988) and is similar to the wall binding domains of other proteins (2.11). This suggests that Epr is a cell wall bound protease which is localized at this site by specific wall binding motifs and also by electrostatic interactions with anionic polymers. However, the interaction of Epr with the cell wall has not been demonstrated.

2.15.6 The *nprB* gene

The *nprB* gene encodes a minor extracellular neutral protease with a molecular mass of 37 kDa which is 56% homologous to the major neutral protease of *B. subtilis*, NprE (Tran *et al.*, 1991). This protease is synthesized as a preproenzyme with a prepro-region composed of 223 amino acid residues, similar to the prepro-regions found in other *B. subtilis* extracellular proteases.

2.15.7 The *vpr* gene

The final extracellular minor protease identified to date is the product of the *vpr* gene. The Vpr protease has a molecular mass of 28 kDa and is completely inhibited by PMSF but not EDTA, suggesting that it is a serine protease (Sloma *et al.*, 1991). Vpr is synthesized as a primary product of 806 amino acid residues with a mature protein of approximately 68 kDa. Since the estimated molecular mass of Vpr is only 28 kDa this protein must also undergo processing.

2.15.8 Summary

Amino acid alignments of the extracellular serine proteases AprE, Bpf, Epr and Vpr indicate that they share homology across the entire sequence, ranging from 25% to 40% (Sloma *et al.*, 1988). Not surprisingly, the most conserved regions are those close to the active site of AprE. The two neutral metalloproteases of *B. subtilis*, NrpE and NprB, show a significant degree of homology (56%) (Tran *et al.*, 1991). Finally, the Mpr protease from *B. subtilis* exhibits almost no homology to bacterial proteases and is more similar to proteases of eukaryotic origin. The latter observation only serves to confuse the evolutionary origins of Mpr.

2.16 Strains of *B. subtilis* deficient in extracellular proteases

It is possible that the production of large quantities of extracellular proteases by *B. subtilis* is a potential factor which limits the use of this organism for the large scale production of heterologous proteins. However, the identification of a large number of genes encoding extracellular proteases has opened up the possibility of isolating strains in which some, or all, of these genes

have been inactivated. The resulting strains may prove valuable as hosts for heterologous protein production since little or no proteolytic degradation of the product would be expected.

A range of protease deficient *B. subtilis* strains have now been constructed by several investigators, each with differing numbers of inactivated proteases. Recently Wu *et al* (1991) have reported the construction of a *B. subtilis* strain, WB600, which is deficient in six extracellular proteases. When compared to the protease proficient wild type strain, *B. subtilis* WB600 (*aprE*, *nprE*, *bpf*, *mpr*, *epr* and *nprB*) displays only 0.32% of the extracellular protease activity.

The stability of proteins in the culture supernatant of *B. subtilis* WB600 was investigated by monitoring the degradation of secreted β-lactamase in culture supernatants. In the wild type strain (*B. subtilis* 168) the half life of secreted β-lactamase in the culture supernatant was 1.5 hours which could be increased to 41 hours when WB600 was used as the host. Therefore, the initial evidence suggests that using protease deficient strains as hosts provides a possible route to improving product recovery. However, the use of such strains as production hosts may also have severe limitations (2.17).

2.17 The influence of extracellular proteases on the activity of autolysins

It has been postulated that extracellular proteases are involved in the control of autolytic activity in *B. subtilis* (Jolliffe *et al.*, 1980). This hypothesis was developed from studies investigating the turnover of peptidoglycan in *B. subtilis* strains with an increased or decreased ability to produce extracellular proteases. The studies revealed that the rate of turnover of peptidoglycan was sensitive to the presence of extracellular protease in the culture supernatant. Mutants deficient in the synthesis of extracellular proteases exhibited an increase in the rate of peptidoglycan turnover as autolytic activity proceeded unchecked. Conversely, mutant strains which hyperproduce extracellular proteases were shown to have a decreased rate of peptidoglycan turnover as the modulating affect of the proteases was increased.

The hypothesis is further supported by studies on *B. subtilis* DB104 which is deficient in the major extracellular proteases, AprE and NprE. Strain DB104 has an increased tendency to lyse towards the end of exponential growth when compared to a protease proficient strain which is indicative of increased autolytic activity in the absence of proteases (Coxon, 1990; Coxon *et al.*, 1991). It seems that at least some extracellular proteases are involved in the control of the autolysins. In addition to deficiencies in extracellular proteases, the absence of an

energized cytoplasmic membrane can also lead to increased cellular autolysis (Jolliffe *et al.*, 1981).

The tendency of protease deficient strains to undergo increased autolysis could prove as problematic for the production of heterologous proteins as the presence of the proteases themselves, since cellular lysis is accompanied by the liberation of intracellular proteases (and other intracellular molecules). In such a situation the protein of interest is likely to be degraded by released intracellular proteases and/or be contaminated by cellular debris to such an extent that considerable downstream processing of the product would be required. Both of these consequences would be unfavourable for cost-effective production of proteins. The addition of protease inhibitors such as EDTA and PMSF to the growing cultures is not a solution since they affect the growth and physiology of the cell..

2.18 α-Amylase, an enzyme of industrial importance

The genus *Bacillus* is a major source of enzymes for industrial applications. The extensive use of *Bacillus* species for this purpose is due to a high secretion capacity and also to the non-pathogenic nature of most members of the genus. The enzymes of interest are almost exclusively extracellular and degradative; their biological function is to degrade polymeric macromolecules in the immediate vacinity of the bacterium for utilization as alternative growth substrates. An enzyme of major industrial importance is α-amylase which randomly hydrolyzes internal 1, 4-α-bonds within starch and glycogen resulting in a rapid reduction in the viscosity of the substrate (termed liquefaction). The products of the reaction are malto-oligosaccharides which are two to six units long, depending on the nature of the substrate.

In general, α-amylases from *Bacillus* species and other bacteria are secreted into the culture supernatant at the end of exponential growth and are commonly subjected to temporal regulation and catabolite repression. An additional cell-bound α-amylase has been detected in *B. subtilis* which is immunologically related to the secreted form from the same organism (Haddaoui *et al.*, 1995). Of all the α-amylases isolated from *Bacillus* species the enzymes from *B. licheniformis* (AmyL), *B. amyloliquefaciens* (AmyQ) and *B. stearothermophilus* (AmyS) are the most homologous and the secreted *B. subtilis* enzyme (AmyE) is less related to the other three. At the amino acid level, the mature AmyL protein is 80% and 63% homologous to AmyQ and AmyS, respectively. The most striking difference between the three related enzymes is

that AmyS possesses a carboxyl terminal extension of 32 amino acid residues (Gray *et al.*, 1986).

It should be noted that *B. licheniformis* produces an additional extracellular α-amylase which differs from AmyL in amino acid sequence and catalytic activity (Kim *et al.*, 1992). The conserved regions of this enzyme exhibit only limited similarity to AmyL and the enzymes from other *Bacillus* species. This additional α-amylase of *B. licheniformis* will not be considered further as this study is primarily concerned with AmyL, an enzyme commonly used for industrial starch processing.

Industrially, liquefying α-amylases are produced from *B. amyloliquefaciens* and *B. licheniformis* and reduce starch to maltohexaose or maltopentaose, respectively. A major use of the α-amylases from *B. licheniformis* and *B. amyloliquefaciens* is the liquefaction of starch prior to saccharification during the production of sugar syrups (Fogarty, 1983). The enzyme from *B. licheniformis* is highly thermostable (2.18.1) and has a temperature optimum of 90°C to 95°C which permits starch processing at elevated temperatures. This property alone enables faster, more efficient, and therefore, more cost-effective starch liquefaction. For these reasons the enzyme from *B. licheniformis* is generally used in preference to the enzyme from *B. amyloliquefaciens* for industrial starch processing. In addition, α-amylases are widely utilized by the detergent, baking and brewing industries for various starch hydrolysing processes.

2.18.1 The α-amylase from *B. licheniformis* is highly thermostable

The product of the *B. licheniformis amyL* gene is a precursor protein of 512 amino acid residues. The first 29 residues of the precursor constitute a typical Gram-positive N-terminal signal peptide with a cluster of four positively charged amino acids followed by an extensive hydrophobic domain (Stephens *et al.*, 1984). During translocation, the AmyL precursor is processed by the signal peptidase at a consensus cleavage site to yield the enzymatically active mature protein which is 483 amino acids long (Gray *et al.*, 1986). The increased thermal stability of this enzyme compared with AmyQ and AmyS has been demonstrated by measuring the half life of the enzymes at 90°C. Under these conditions, the half lives for the enzymes were 270, 50 and 2 min for AmyL, AmyS or AmyQ, respectively (Tomazic and Klibanov, 1988). Because of the commercial importance of AmyL, attention has focused on determining the basis of this thermostability and several studies have identified specific regions of the protein which may be involved.

Tomazic and Klibanov (1988) investigated the reasons for the different irreversible thermoinactivation profiles of the related *Bacillus* α-amylases. The increased stability of thermophillic enzymes at elevated temperatures can often be attributed to the formation of salt bridges which stabilize reversible thermal unfolding within the protein (Perutz, 1978). In the presence of concentrated inorganic salts which weaken such electrostatic interactions, the thermostability of AmyL was reduced to a level comparable with that of AmyQ. The thermostability of AmyQ was also affected by the presence of concentrated inorganic salts but the relative affect was not as pronounced as that for AmyL, suggesting that salt bridges are important in mediating thermostability of this enzyme. Three lysine residues, lysine-88, lysine-253 and lysine-385, were identified within AmyL which are not present in related α-amylases and therefore, may be involved in stabilizing salt bridge formation (Tomazic and Klibanov, 1988).

In addition to these lysine residues, two regions of AmyL, spanning amino acids 177 to 186 (region I) and 255 to 270 (region II), are also thought to be involved in thermostability (Suzuki *et al.*, 1989). Region I seems to play a major role in thermostability while region II seems to play a lesser role. Both regions affect thermostability independently and additively. Interestingly, alteration of amino acids within AmyQ which made corresponding domains more like regions I and II from AmyL increased the thermostability of chimeric enzymes to levels almost as high as that of AmyL, confirming the involvement of these regions, and in particular charged amino acids, in the resistance to irreversible thermal inactivation (Suzuki *et al.*, 1989).

When compared with the primary structures of AmyQ and AmyS, a prominent feature of AmyL is the presence of a larger number of histidine residues; with 14, 11 and 24 histidines in each protein, respectively (Declerck *et al.*, 1990). Since histidine is one of the less commonly found amino acids, Declerck *et al* (1990) postulated that the presence of such a large number of histidine residues in the most thermostable of the three related *Bacillus* α-amylases is likely to contribute to the stability of this protein at elevated temperatures. Using amber supressors, a histidine residue important to the thermostability of AmyL was identified at position 133 and amino acid substitutions at this position can either increase or decrease thermostability depending on the amino acid inserted.

The irreversible thermal inactivation of AmyQ and AmyS occurs by a monomolecular conformational scrambling type mechanism (Tomazic and Klibanov, 1988). It appears that the thermal inactivation of AmyL which does

occur, albeit at a lower level, is mediated by the deamidation of asparagine and/or glutamine residues within the protein and not by monomolecular scrambling. The failure of different biochemical and genetical studies to identify common domains involved in AmyL thermostability suggests that several different regions of the protein are important.

2.18.2 The role of calcium in the stabilization of α-amylases

α-Amylases, whether from prokaryotic or eukaryotic origin, display a common requirement for divalent metal ions. Historically, the stabilizing effect of metal ions on α-amylase was observed by Wallerstein more than eighty years ago and was subsequently patented (Wallerstein, 1909). Of all the commonly occurring divalent metal ions, only calcium and zinc can be detected in complex with α-amylases in stoichiometrically significant quantities (Vallee *et al.*, 1959). In the presence of divalent metal ions, α-amylases are generally resistant to the activity of proteolytic enzymes whilst α-amylase preparations from which divalent metal ions have been removed become highly vulnerable to proteolysis (Stein and Fischer, 1958; Machius *et al.*, 1995). Vallee *et al* (1959) have shown this effect to be mediated solely by the presence or absence of calcium, with every enzyme molecule complexing at least one Ca^{2+} ion.

In the presence of chelating agents, such as EDTA, pure preparations of AmyL can be reversibly inactivated as Ca^{2+}ions are sequestered from the protein and solvent. The enzymatic activity of AmyL can be subsequently restored by the addition of excess calcium ions, demonstrating the importance of calcium for the activity of α-amylase (Violet and Meunier, 1989). In addition, the presence of calcium plays a key role in the thermostability of AmyL as Violet and Meunier (1989) demonstrated that in the absence of calcium the enzyme becomes more susceptible to irreversible thermal inactivation. This stabilizing effect is calcium-specific as no other metal ion or substrate can affect the activity in a similar manner.

Based on the crystal structure of the porcine pancreatic α-amylase, the role of the essential calcium ion in the structure and function of this enzyme has been postulated (Buisson *et al.*, 1987). Buisson *et al* (1987) observed that the ligands for the calcium ion are located in domain-A and domain-B (2.18.3), and that the active site cleft for this enzyme is located between these two domains. They postulated that the conformation of the active site cleft is stabilized by the presence of a ionic bridge between domains-A and -B which is mediated by the presence of the calcium ion. Due to the structural similarities of α-amylases from

various origins it is likely that the calcium ion assumes a similar role in AmyL and other related amylolytic enzymes.

2.18.3 The crystal structure of AmyL

Throughout evolution the tertiary structures of proteins with similar functions are often more conserved than their primary structures (Chothia and Lesk, 1986). This holds true for α-amylases and other related amylolytic enzymes from different organisms. For example, the Taka-amylase A from *Aspergillus oryzae* and the porcine pancreatic α-amylase, for which the crystal structures have been solved, have only 23% identical amino acids but exhibit very similar three-dimensional structures (Matsuura *et al.*, 1984; Buisson *et al.*, 1987). This allows certain predictions to be made about the structures of α-amylases for which the crystal structure has not yet been solved.

In general, α-amylases contain an α/β-barrel as the central domain-A with a carboxy terminal domain-C, which is separated from domain-A by domain-B. AmyL differs from other enzymes in that domain-B possesses a β-sheet composed of six twisted and loosely connected β-strands which form a barrel with a large cavity in the centre (Machius *et al.*, 1995).

The coordinates of the Taka-amylase A were used to construct a molecular model of the α-amylase from *B. stearothermophilus* (Holm *et al.*, 1990). The enzyme was subjected to random mutagenesis in an attempt to analyse its structure and function. The location of mutations on the model and the corresponding mutant phenotypes revealed, not surprisingly, that amino acid substitutions with detrimental affects on activity were clustered at the active site, along the substrate binding cleft and at the interface between the central α/β barrel and the C-terminal domain. As expected, exposed regions of the protein were more tolerant to substitutions. Because of the relatedness of AmyL and AmyS at both the amino acid and structural levels it is reasonable to assume that AmyL would be affected in a similar manner to substitutions in equivalent amino acids.

The crystal structure of calcium-depleted AmyL has recently been solved (Machius *et al.*, 1995). Domain-A is made up of an α/β-barrel with complex loop structures that connect the strands and helices making up the barrel. The loop regions on the C-terminal side of the barrel, which connect α-helices to β-strands, are more complex than those on the N-terminal side and contain the active site of the enzyme and also the main calcium binding site. Domain-B is complex and consists of six β-strands which constitute a loosely connected and twisted antiparallel β-sheet. Two larger two-stranded sheets fold back onto the

antiparallel β-sheet forming a barrel-type structure with a large cavity in the centre. Of the three domains of AmyL, domain-B differs most from the corresponding domain of other α-amylases. Finally, no function has been currently assigned to domain-C although it is known to contain a "greek key" type motif. Domain-C from AmyS has been proposed to be involved in starch hydrolysis by orientating the active site cleft of domain-A with the amylose chain (Holm *et al.*, 1990). This is also likely to be the case for AmyL.

The active site of AmyL is contained within domain-A and the three catalytically important residues are conserved amongst the related *Bacillus* α-amylases, and also those from other organisms. The three residues involved in the catalytic activity of AmyL are aspartate-231, glutamate-261 and aspartate-328. In addition to the three residues making up the active site, other regions around the active site are also well conserved and may be important in substrate binding.

The structure of AmyL was solved using crystals obtained in the absence of calcium and as a consequence the calcium ion could not be placed in the structure. However, since the calcium binding site is well conserved throughout α-amylases, its position was determined using the data obtained from other enzymes. The main calcium binding site of AmyL is situated between the C-terminal end of the central β-barrel of domain-A and domain-B and connects the two regions. It is considered to be in the form of a distorted pentagonal bipyramid similar to that reported for the α-amylase from *Aspergillus niger* (Boel *et al.*, 1990) with three strictly conserved amino acid residues directly involved in calcium binding; namely asparagine-104, aspartate-200 and histidine-235.

2.18.4 New perspectives on the thermostability of AmyL

With the elucidation of the crystal structure of AmyL tentative conclusions can be made regarding the previously described theories attempting to account for the enhanced thermostability. The structure has revealed that the lysine residues (88, 253 and 385) initially thought to form stabilizing salt bridges do not participate in any form of ionic interaction. The structure of AmyL is stabilized by ionic interactions but the only salt bridges at the interface between domains occur between aspartate-60 and arginine-146, and also between aspartate-204 and lysine-237, therefore, ruling out the lysines proposed by Tomazic and Klibanov (1988).

Region I of AmyL exists as a loop on the surface of domain-B. The corresponding region of AmyQ is two amino acid residues longer and this increase in size of the looped region may lead to increased mobility at elevated temperatures, causing a consequent decrease in the thermostability of AmyQ.

However, determination of the crystal structure of AmyL did not allow the identification of any specific interactions in regions I and II which may be responsible for thermostability. This comparative approach is hindered by the lack of accurate structural data for the least thermostable of the related *Bacillus* α-amylases, AmyQ.

2.18.5 Summary

α-Amylases are used for large scale starch hydrolysis in numerous industrial processes and therefore, the sale of such enzymes can often represent a considerable proportion of the revenue of companies which produce enzymes on an industrial scale. As a consequence, considerable effort has been invested in determining the properties and structure of relevant α-amylases. Not surprisingly, significant attention has been focused on the thermostable α-amylase from *B. licheniformis*. The crystal structures of α-amylases from prokaryotic and eukaryotic origins are now known and it is clear that the proteins are well conserved at the structural level. Consequently, using this information, predictions can be made about the structure and function of other α-amylases which have not been extensively characterized.

2.19 The *B. subtilis xyl* regulon

The genes for the utilization of xylose polymers in *B. subtilis* 168 and W23 have been identified and cloned (Hastrup, 1988; Wilhelm and Hollenberg, 1984). The genes are organized into a regulon which is inducible by xylose. Expression of the genes of the regulon is negatively regulated at the level of transcription (Hastrup, 1988; Gartner *et al.*, 1988) and is subject to catabolite repression (CR) (Jacob *et al.*, 1991; Kraus *et al.*, 1994).

The *B. subtilis* 168 *xyl* regulon, indicated in Fig. 2.7, is composed of five genes; *xynC*, *xynB*, *xylR*, *xylA* and *xylB*. *xynB* encodes β-xylosidase, *xylR* the xylose repressor, *xylA* xylose isomerase and *xylB* xylulokinase. The function of *xynC* is unclear although its role as xylose permease has been postulated due to the presence of potential membrane spanning domains (Hastrup, 1988).

The genes of the regulon are divided into three transcriptional units; *xynC-xynB*, *xylR* and *xylA-xylB*. *xynC-xynB* and *xylA-xylB* are transcribed from two promoter/operators termed P_1O and P_2O, respectively, while *xylR* is transcribed from P_R. The operator sites of the *xyl* regulon of *B. subtilis* 168 and W23 are 25 bp in length, composed of two 10 bp IR sequences separated by 5 non-IR bases and are situated 10 bp downstream from their respective promoters (Fig. 2.7). The *xyl* operator regions of 168 and W23 are identical except for the 5 non-IR

bases. Mutational analysis has shown that the 5 bp sequence separating the IR's is involved in sequence-specific operator recognition (Gartner *et al.*, 1992). In addition, Kreuzer *et al* (1989) have suggested that the second, outermost T of the IR sequence in *B. subtilis* W23 is involved in sequence-specific recognition of the operator by the repressor.

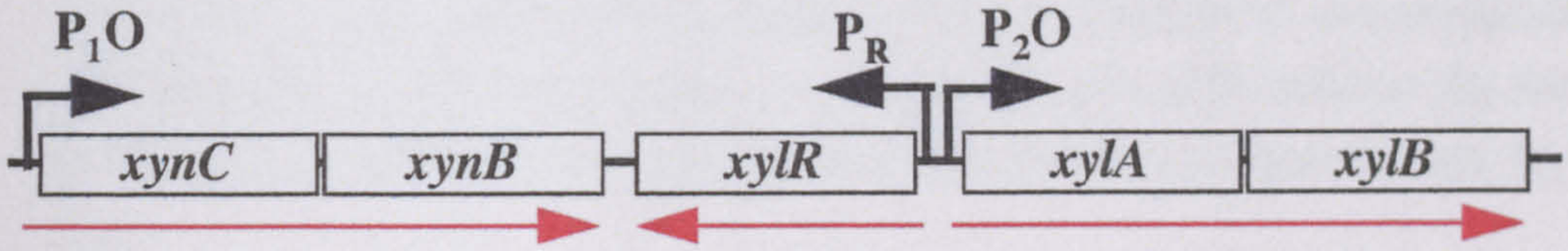
Control of the regulon is mediated by the *trans*-acting XylR repressor which, in the absence of xylose, binds to the operator sites and represses initiation of transcription from P_1 and P_2 (Gartner *et al.*, 1988; Kreuzer *et al.*, 1989; Gartner *et al.*, 1992). The presence of xylose in the growth medium leads to the inactivation of the cf XylR. In common with many other DNA binding proteins, XylR has an α-helix-turn-α-helix DNA binding motif (Dodd and Egan, 1990) and it has been suggested that XylR interacts with the operators over two full helical turns (Gartner *et al.*, 1992).

CR is a global regulatory mechanism which causes the repression of certain genes in the presence of rapidly metabolizable carbon sources. CR in *B. subtilis* and other Gram-positives has been reviewed extensively (Stewart, 1993; Hueck *et al.*, 1994; Hueck and Hillen, 1995) and is negatively regulated, at least in part, by the action of the catabolite control protein (CcpA). CcpA interacts with catabolite responsive elements (CRE's) in the DNA of CR-sensitive genes preventing transcription. A component of the phosphoenolpyruvate-dependent phosphotransferase system (PTS), Hpr, is also involved. The consensus CRE is 14 bp long (Weickert and Chambliss, 1990) and may be located within the coding regions of genes subject to CR or more commonly, close to their promoters (Hueck *et al.*, 1994).

The expression of *xylA-xylB*, has been shown to be subject to CR (Jacob *et al.*, 1991; Kraus *et al.*, 1994) and is mediated by a CRE which conforms to the consensus CRE common to other catabolite repressible genes (Weickert and Chambliss, 1990; Hueck *et al.*, 1994). The CRE for *xylA-xylB*, 5'-TGGAAGCGTAAACA-3', is located 139 bases downstream from the P_2 -10 region within the *xylA* structural gene. Mutations in this CRE making it more like the consensus lead to an increase in CR (Kraus *et al.*, 1994). In addition to CRE, it is possible that the CR of inducible genes, such as *xylA*, is mediated by an inducer exclusion type mechanism when glucose is used as the carbon source since glucose repression is reduced when *xylR* is inactivated (Kraus *et al.*, 1994). In this system inducer exclusion may act at the level of xylose uptake or alternatively, glucose may compete with xylose for binding to XylR.

The xylose inducible control elements have been used by Novo Nordisk (Bagsværd, Denmark) to construct expression plasmids to investigate the

secretion of α-amylases in *B. subtilis*. Plasmid pSX63 (Appendix 5) is based on pUB110 and encodes a copy of the *amyL* gene expressed from $P_1$0. To achieve inducibility, the *xylR* gene is also present on the plasmid to provide the *trans*-acting repressor protein. pSX63 also includes a chloramphenicol acetyl transferase (*cat*) gene to facilitate the selection of plasmid-containing transformants. Therefore, strains of *B. subtilis* harbouring pSX63 synthesize α-amylase only in the presence of xylose. This plasmid was used in this study as the starting plasmid for the construction of vectors to facilitate the investigation of protein secretion in *B. subtilis* (4.1).



P_1O

-35 **-10**

5'- AAAAAGATG **TTGAAA** AAGTCGAAAGGATTTTA **TAATAT** TAAGT

C*AAG **TTAGTTTGTT** TGATC **AACAAACTAA** T -3' 292 bp to ATG of *xynC*

P_2O

-35 **-10**

5'- AAAAAAAATA **TTGAAA** ATACTGACGAGGTTATA **TAAGAT** GAAAA

T*AAG **TTAGTTTGTT** TAAAC **AACAAACTAA** T -3' 72 bp to ATG of *xylA*

Fig. 2.7. Schematic representation of the *B. subtilis* 168 *xyl* regulon and DNA sequences of the promoter/operator elements.

The direction of transcription of the relevant genes is indicated by red arrows. The promoters are indicated in red and the operator sequences in blue. The preferred mRNA start is after the asterix.

2.20 Project aims

B. subtilis and other members of the genus secrete proteins directly into the culture supernatant efficiently and at a high level. This fact, combined with the GRAS status of most *Bacillus* species, has meant that these bacteria have been used extensively for the industrial scale production of commercially important (homologous) proteins. However, the use of *B. subtilis* for the production of heterologous proteins has been limited, at least in part, by the cell wall.

The physico-chemical properties of the cell wall are likely to influence the passage of secretory proteins across the wall cylinder following translocation, as are the properties of exoproteins. As part of the European Commission Biotechnology programme 1992-1994 the Newcastle Molecular Microbiology Group, in collaboration with the Bacterial Gene Technology Group at Novo Nordisk, was contracted to determine the extent to which the *B. subtilis* cell wall acts as a barrier to the secretion of exoproteins.

Therefore, within the framework of the programme, the aim of this study was to engineer variants of a well secreted model protein, α-amylase, and use them to determine the effect on post-translocational secretion events and in particular, those relating to exoprotein folding and the interaction of exoproteins with the *B. subtilis* cell wall.

CHAPTER 3

MATERIALS AND METHODS

3.1 Growth and maintenance of bacterial strains

The *B. subtilis* and *E. coli* strains used during this study are shown in Table 3.1. Unless stated otherwise, the strains were maintained on penassay agar plates (Appendix 3) containing antibiotics where required. For long-term storage, cultures were grown overnight in 2xYT broth and stored at -70°C after the addition of glycerol to 15% v/v.

Batch cultivation was carried out at 37°C in a Gallenkamp orbital incubator with shaking at 180 rpm. In both solid and liquid media, unless stated otherwise, chloramphenicol was used at 6 μg/ml and ampicillin at 100 μg/ml. Xylose (1% w/v) was also included as required to induce the synthesis of α-amylase.

Table 3.1. Bacterial strains used during this study

Strain	Genotype	Source
B. subtilis DN1885	*amyE*	Novo Nordisk
E. coli SJ2	*lacI*^{q}*ZΔM15*, r$^-$ m$^+$	Novo Nordisk
E. coli XL1-Blue	*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac*, [F' *proAB*, *lacI*^{q}*ZΔM15*, *Tn10*(tet^r)]	Stratagene

3.2 Plasmids

The plasmid vectors used in this study are shown in Table 3.2. For clarity related plasmids derived during this work are not included.

Table 3.2. Plasmids used during this study

Plasmid	Host	Selective marker	Source
pSX63	*B. subtilis*	Cm	Novo Nordisk
pKS301	*E. coli*/*B. subtilis*	Ap, Cm	This study
pKS401	*E. coli*/*B. subtilis**	Ap, Cm	This study
pUC19	*E. coli*	Ap	New England Biolabs
pBluescript SK$^-$	*E. coli*	Ap	Stratagene

* By integration

3.3 Precipitation of DNA

DNA was precipitated in 1.5 ml tubes by the addition of 0.1 volumes of sodium acetate (3 M, pH 4.8) followed by 2.5 volumes of pre-cooled ethanol or 1 volume of pre-cooled isopropanol. The tubes were incubated at -20°C for 10 min to aid the precipitation, after which the DNA was pelleted by microcentrifugation (13 krpm, 15 min, room temperature). The supernatant was aspirated and the DNA pellet washed with 500 µl of 70% v/v ethanol to remove residual salt. Following a second microcentrifugation step, the pelleted DNA was dried in a vacuum dessicator and resuspended in a suitable volume of sterile distilled water.

3.4 Purification of plasmid DNA

3.4.1 Small scale extraction of plasmid DNA - STET method

This rapid and convenient method of plasmid extraction from *E. coli* is based on the method of Holmes and Quigley (1981). The plasmid DNA isolated was only suitable for restriction analysis. The cells from 1.5 ml of an overnight culture in penassay broth, containing relevant antibiotics, were harvested by microcentrifugation (13 krpm, 5 min, room temperature) and resuspended in 200 µl of STET buffer (Appendix 3).

The cell suspension was boiled for 1 min and allowed to cool to room temperature. Cellular debris was pelleted by microcentrifugation (13 krpm, 10 min, room temperature) and removed using a toothpick. Contaminating RNA was then digested by the addition of 1 µl of a 10 mg/ml RNase A solution (DNase free) followed by incubation at 37°C for 15 min.

Plasmid DNA was precipitated by the addition of 200 µl of pre-cooled isopropanol and mixing by inversion. Precipitated DNA was pelleted by microcentrifugation (13 krpm, 15 min, room temperature) and the supernatant aspirated. The DNA pellet was dried using a vacuum dessicator and resuspended in 50 µl of sterile distilled water. This method was used to screen large numbers of transformants and to detect the presence of inserts of the correct size by restriction analysis (3.5).

3.4.2 Small scale extraction of plasmid DNA - acid phenol method

This method is based on that described by Kieser (1984) and can be used to extract plasmid DNA from *E. coli* and *B. subtilis*. The resultant plasmid DNA is suitable for restriction analysis, cloning, sequencing and PCR. The cells from 2 ml of an overnight culture in penassay broth were harvested by microcentrifugation (13 krpm, 5 min, room temperature) and resuspended in 500 µl of TES (Appendix 3) containing 2 mg/ml lysozyme. For *B. subtilis*, the cell

suspensions were incubated at 37°C for 30 min and for *E. coli* at room temperature for 15 min.

Following incubation, 250 µl of 0.2 M NaOH/1% w/v sodium dodecyl sulphate (SDS) was added and mixed immediately by vortexing. The sample was incubated at 65°C for 20 min, allowed to cool to room temperature and then 200 µl of acid phenol was added. The phases were mixed by vortexing and microcentrifuged (13 krpm, 5 min, room temperature). The upper phase containing plasmid DNA was removed and transferred to a new tube. The DNA was precipitated by the addition of 70 µl of 3 M sodium acetate and 700 µl of isopropanol. The solutions were mixed by inversion and the precipitated DNA pelleted by microcentrifugation (13 krpm, 15 min, room temperature).

The DNA was resuspended in 50 µl of distilled water and extracted again by adding 10 µl of 3 M sodium acetate and 25 µl of neutral phenol (Appendix 3) as described above. Following microcentrifugation, the upper phase was transferred to a new tube and DNA precipitated with 55 µl of isopropanol. The DNA was pelleted by microcentrifugation and washed twice with 400 µl of 70% v/v ethanol. The DNA pellet was dried using a vacuum dessicator and resuspended in 50 µl of sterile distilled water and 1 µl of a 10 mg/ml solution of RNase A (DNase free). The sample was incubated at 37°C for 15 min to allow degradation of any contaminating RNA and then stored at -20°C.

3.4.3 Extraction of plasmid DNA using the Qiagen-tip-100 kit

Qiagen-tip-100 columns (Qiagen) contain an immobilized ion-exchange resin to which plasmid DNA from cell lysates can be bound, washed and subsequently eluted. The columns can be used to extract plasmid DNA from *E. coli* and *B. subtilis*. The plasmid DNA isolated using these columns was suitable for restriction analysis, cloning, sequencing and PCR. This kit was used to prepare up to 100 µg of plasmid DNA from *E. coli* and *B. subtilis*. The cells from 100 ml of an overnight penassay broth culture were harvested by centrifugation (10000 *g*, 15 min, 4°C) and resuspended in 4 ml of P1 buffer (Appendix 3). For the isolation of plasmid DNA from *B. subtilis*, lysozyme was included in P1 buffer at a final concentration of 5 mg/ml and the cell suspensions incubated at 37°C for 15 min. *E. coli* was not pre-treated with lysozyme. P2 buffer (4 ml, Appendix 3) was then added to the suspension followed by incubation at room temperature for 5 min. Pre-chilled P3 buffer (4 ml, Appendix 3) was added and the mixture was incubated on ice for 15 min to precipitate denatured proteins and chromosomal DNA.

Insoluble material was pelleted by centrifugation (10000 *g*, 30 min, 4°C) and the supernatant transferred to a tip-100 column which had been equilibrated

with 4 ml of QBT buffer (Appendix 3). The bound DNA was washed twice with 10 ml of QC buffer, eluted with 5 ml of QF buffer, precipitated (3.3) and then resuspended in the required volume of distilled water, which was typically 250 µl.

3.5 Digestion of DNA with restriction endonucleases

All restriction enzymes used in this study were obtained from NBL Gene Sciences (Cramlington, UK). Restriction digests were carried out at 37°C in accordance with the manufactures instructions and using the buffers provided with the enzymes. For small scale restriction analysis (1-4 µg of DNA) the final volume of the reaction was typically 20 µl and contained 1-3 units of each enzyme. Large scale digestions (up to 25 µg of DNA) were carried out in a final volume of 100 µl with approximately 20 units of each restriction enzyme.

3.6 Electrophoresis of DNA

Prior to gel loading, DNA samples were mixed with 0.1 volumes of x10 DNA loading buffer (Appendix 3). All agarose gels were cast in x0.5 TBE buffer (Appendix 3) and electrophoresed in a DNA sub-cell (Bio-Rad) using the same buffer. DNA fragments above or below 1000 bp were electrophoresed on 1% or 2% w/v agarose gels, respectively. Gels were stained with a 0.5 µg/ml solution of ethidium bromide, viewed by UV transillumination and photographed using either Polaroid 667 film or a gel documentation system (Bio-Rad).

3.7 Purification of DNA from agarose gel slices using the Qiaquick kit

The Qiaquick kit (Qiagen) utilizes the same DNA binding resin as the Qiagen-tip-100 columns (3.4.3) to purify DNA from agarose gel slices following digestion (3.5) or amplification with PCR (3.12). The DNA to be purified was electrophoresed (3.6) and the fragment of interest excised from the gel using a clean scalpel and transferred to a 2 ml microcentrifuge tube. QX1 buffer was added to the agarose slice to a final volume of 2 ml and the tube was incubated at 50°C until the agarose was completely solubilized (typically 15 min). The solubilization process was helped by occasional mixing of the tube contents.

The gel/DNA solution was loaded onto a Qiaquick column inside a new tube followed by microcentrifugation (13 krpm, 1 min, room temperature) to bind the DNA to the column. The flow-through fraction was removed from the tube and the loading procedure repeated until all the gel/DNA solution had been passed through the column. The DNA bound to the column was washed with 750 µl of PE buffer followed by microcentrifugation as above. The flow-through fraction was removed and the column microcentrifuged to remove any residual

wash buffer. DNA was eluted from the column into a fresh tube by the addition of distilled water (50-100 μl) followed by microcentrifugation. If required, the DNA fragment was precipitated (3.3) and resuspended in a smaller volume of distilled water.

3.8 Ligation of DNA fragments

DNA fragments were ligated following restriction digestion (3.5) and purification (3.7). Vector and insert DNA were combined in a tube in an approximate molar ratio of 1:3 (as estimated by agarose electrophoresis) together with x2 ligation buffer (Appendix 3) and an appropriate amount of T_4 DNA ligase (NBL Gene Sciences). The final volume of the ligation reaction was typically 30 μl. The tube was incubated at 4°C overnight and then used to transform *E. coli* or *B. subtilis*.

3.9 Transformation of *B. subtilis* using natural competence

During transition from exponential to stationary phase, *B. subtilis* induces a number of developmental processes, one of which is the ability to take up exogenous DNA. The development of competence has been exploited for the introduction of plasmid or chromosomal DNA into laboratory strains of *B. subtilis*.

3.9.1 Preparation of naturally competent *B. subtilis* cells

Several colonies of *B. subtilis* from a fresh agar plate were resuspended in 2 ml of KM-1 medium (Appendix 3). A volume of this suspension was used to inoculate 30 ml of KM-1 medium to an optical density at 450 nm (OD_{450}) of 0.02. The culture was incubated at 37°C with shaking at 180 rpm until stationary phase.

At t_1, 3 ml of the KM-1 culture was used to inoculate 30 ml of KM-2 medium (Appendix 3) and this culture was incubated for a further 45 min. The culture was then transferred to a sterile disposable centrifuge tube and the cells were harvested by centrifugation (5000 *g*, 15 min, room temperature). The supernatant was removed and the cells were resuspended in 2.4 ml of the supernatant and 600 μl of sterile 86% v/v glycerol. Aliquots (50 μl) of the cells were flash frozen in a dry ice/ethanol bath and stored at -70°C until required.

3.9.2 Transformation of naturally competent *B. subtilis*

BTF solution (Appendix 3) was freshly prepared and pre-warmed at 42°C. An aliquot of frozen competent cells was thawed and incubated at 42°C for 3 min. The BTF solution (50 μl) was then mixed with 50 μl of cells and 1-10 μl

(up to 10 μg) of DNA. The transformation mixture was incubated at 37°C for 20 min with gentle shaking to allow phenotypic expression. The cells were then plated onto penassay agar containing the required antibiotics and incubated at 37°C for 12-36 hours.

3.10 Transformation of *E. coli*

3.10.1 Preparation of *E. coli* cells for electroporation

The cells from a single colony on a penassay agar plate were used to inoculate 20 ml of L-broth (Appendix 3) and the culture was grown overnight at 37°C with shaking. A volume of the culture was used to inoculate 200 ml of pre-warmed L-broth to an OD_{600} of 0.06. The culture was incubated as before until the OD_{600} reached 0.8 to 1.0.

The flask containing the cells was chilled on ice for 30 min and the cells harvested by centrifugation (10000 *g*, 15 min, 4°C) and washed successively with 200 ml and 100 ml aliquots of pre-cooled sterile distilled water. Finally, the cells were washed with 4 ml of sterile 10% v/v glycerol and then resuspended in 10% v/v glycerol to final volume of 600 μl. The cells were flash frozen in 50 μl aliquots using a dry ice/ethanol bath and stored at -70°C until required.

3.10.2 Electroporation

An aliquot of electrocompetent cells was thawed at room temperature and then stored on ice. An appropriate amount of DNA, typically 1-2 μl, was added to 50 μl of cells and the mixture incubated on ice for 1-2 min. The mixture was transferred to a cooled 0.2 cm electroporation cuvette and pulsed (Bio-Rad Gene Pulser) at 25 μF capacitance, 2.5 kV and 200 Ω resistance. These settings should produce a time constant of 4.5-5.0 ms

Immediately after the pulse, 1 ml of SOC (Appendix 3) was added to the cuvette and the cells resuspended. The cell suspension was transferred to a reaction tube and incubated at 37°C with shaking at 180 rpm for 60 min to allow phenotypic expression of antibiotic resistance markers. Following incubation, 100 μl was plated onto penassay agar containing the required antibiotics and incubated at 37°C overnight.

3.10.3 Preparation of chemically competent *E. coli*

A single colony of *E. coli* from a fresh penassay agar plate was used to inoculate 20 ml of L-broth and the culture was grown at 37°C with shaking until the OD_{600} reached approximately 0.8. The culture was then added to 100 ml of pre-warmed L-broth and incubation continued. When the OD_{600} again reached 0.8, 400 ml of pre-warmed L-broth was added and the culture was allowed to

grow to an OD_{600} of 0.6, at which time the flask was removed from the incubator and chilled in an ice water bath for 15 min.

The cells were harvested by centrifugation at (5000 *g*, 20 min, 4°C) and gently resuspended in 100 ml of pre-cooled TFBI (Appendix 3). The cells were harvested as before, resuspended in 20 ml of TFBII (Appendix 3) and flash frozen in 100 µl aliquots using a dry ice/ethanol bath. Aliquots of cells were stored at
-70°C until required.

3.10.4 Transformation of chemically competent *E. coli*

Plasmid DNA (1-2 µl) or ligation reaction (10-15 µl) was added to 100 µl of thawed competent cells and the mixture left on ice for 30 min. To facilitate the uptake of DNA, the cells were heat shocked at 42°C for 1.5 min and then transferred to a fresh tube containing 1 ml of penassay broth. The tube was incubated at 37°C for 60 min with shaking to allow phenotypic expression and then 100 µl was plated onto penassay agar containing the required antibiotics. The plates were incubated overnight at 37°C.

3.11 Screening for recombinant plasmids using α-complementation

Certain *E. coli* plasmid vectors, such as pUC19, encode the regulatory regions (*lacI*) and the 146 N-terminal amino acids of the β-galactosidase gene (*lacZ*). An in-frame multiple cloning site is also located within this coding region, and the extra amino acids introduced as a consequence of this do not affect the activity of the complete LacZ protein.

Vectors of this type are commonly used in conjunction with strains, such as *E. coli* XL1-Blue, encoding only the C-terminal part of LacZ which is non-functional on its own. When present together, the plasmid- and host-encoded LacZ fragments associate to form an enzymatically active protein and colonies containing such vectors appear blue on agar plates containing the chromogenic substrate X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) in the presence of the inducer, IPTG (isopropylthio-β-D-galactoside).

This effect is termed α-complementation (Ullmann *et al.*, 1967) and can be used as a screening step to identify recombinant plasmids since, in almost all cases, cloning a DNA fragment into the multiple cloning site inactivates the plasmid-encoded part of LacZ, resulting in the failure to produce an active β-galactosidase. Consequently, colonies harbouring plasmids with inserts appear white on agar plates containing X-Gal and IPTG. α-Complementation was used frequently during this study to identify recombinant plasmids arising after

transformation of *E. coli* by plating onto selective agar plates containing 50 μg/ml X-Gal (in dimethyl formamide) and 0.5 mM IPTG.

3.12 The polymerase chain reaction (PCR)

PCR is an *in vitro* DNA amplification procedure that uses a thermostable DNA polymerase, *Taq*, and two oligonucleotide primers to amplify a specific region of DNA (Mullis and Faloona, 1987). PCR was used in this study to amplify a DNA fragment for the construction of a shuttle vector (4.3) and also to splice specific regions of α-amylase genes during the construction of chimeric α-amylases (5.3). To amplify from plasmid templates the reactions were generally set up as shown in Table 3.3. In some cases it was necessary to set up a negative PCR control which contained all the reaction components except template DNA. This was carried out to ensure that the product amplified during PCR was specific for the template DNA.

Table 3.3. The components of the PCR used to amplify from plasmid templates

Reagent	Volume
dNTP (2.5 mM)	10 μl
Reaction buffer (x10)	10 μl
Primer-1 (20 pmole/μl)	10 μl
Primer-2 (20 pmole/μl)	10 μl
Distilled water	58.5 μl
Taq (5 U/μl)	0.5 μl
Template (0.5 μg/μl)	1.0 μl
Mineral oil	75 μl

During the construction of the chimeric α-amylase genes it was necessary to splice two PCR products which had been synthesized in previous reactions (5.3.3). In this situation the components of the PCR were slightly different from those shown in Table 3.3 because the template DNA for each reaction was composed of two individual PCR products. In general, the gene splicing reactions were set up as shown in Table 3.4.

Continued

Table 3.4. The components of the PCR used for splicing reactions

Reagent	Volume
dNTP (2.5 mM)	10 µl
Reaction buffer (10x)	10 µl
Primer-1 (20 pmole/µl)	10 µl
Primer-2 (20 pmole/µl)	10 µl
Distilled water	49.5 µl
Taq (5 U/µl)	0.5 µl
Template-1 (0.5 µg/µl)	5 µl
Template-2 (0.5 µg/µl)	5 µl
Mineral oil	75 µl

3.12.1 PCR conditions

A Hybaid thermocycler was used for all reactions, and each PCR involved 35 cycles of denaturation, primer annealing and strand extension. In all cases, denaturation was carried out at 92°C for 2 min. This was followed by a primer annealing step of 1 min at the lowest melting temperature of the primer pair, minus 2°C. Strand extension was at 72°C allowing 1 min per 1000 bp of target DNA. After cycle 35, one cycle at 72°C for 10 min was carried out to complete any unfinished strand synthesis. The resultant PCR products were analysed by agarose gel electrophoresis (3.6) and, if required, purified with Qiaquick columns (3.7).

3.13 DNA sequencing

Plasmid DNA, isolated from *E. coli* strains using acid phenol (3.4.2) was sequenced by the dideoxy chain-termination method (Sanger *et al.*, 1977) using the Sequenase kit (Amersham International plc).

3.13.1 Preparation of template DNA

Approximately 5 µg of plasmid DNA (~30 µl of DNA preparation) was made up to a final volume of 50 µl with distilled water and 30 µl of 20% w/v polyethylene glycol-6000/1.5M NaCl was added to precipitate the DNA. The tube was incubated on ice for 1 hour and the DNA pelleted by microcentrifugation (13 krpm, 15 min, room temperature). The supernatant was aspirated, the DNA pellet washed twice with 70% v/v ethanol and the pellet resuspended in 8 µl of distilled water.

3.13.2 Denaturation of template DNA and primer annealing

The double-stranded DNA was denatured by the addition of 2 µl of 2 M NaOH followed by incubation at room temperature for 5 *min*. Following incubation, the DNA was ethanol precipitated (3.3) and resuspended in 7 µl of distilled water. To the denatured template, 1 µl of primer (0.5 pmole/µl) and 2 µl of reaction buffer were added. The tube was incubated at 37°C for 15 *min* to allow primer annealing and then stored on ice until required.

3.13.3 Sequencing reactions and electrophoresis

The labelling and termination reactions were carried out with Sequenase v2 as described in the protocol supplied with the kit and using α-^{35}S-dATP (Amersham International plc) for labelling. Following termination, the reactions were heated to 92°C for 5 min and aliquots (3 µl) electrophoresed at 50 W (constant voltage) on a 6% w/v acrylamide:bis-acrylamide (19:1) gel in x1 TBE buffer. The gel was fixed for 30 min in 5% v/v acetic acid/15% v/v methanol, dried onto 3MM filter paper (Whatman) using a vacuum gel drier (Bio-Rad) and then exposed to X-ray film.

3.14 Quantitation of α-amylase activity

α-Amylase activity was measured using the Phadebas assay kit (Pharmacia Diagnosics). The substrate for the kit is a water insoluble cross-linked starch polymer carrying a blue dye. It is hydrolysed by α-amylase to form blue, water soluble, fragments. The absorbance at 620 nm (A_{620}) of the blue solution is a function of the α-amylase activity in the sample, the latter being determined from the standard curve provided with the kit. The assay was used to determine the α-amylase activity in *B. subtilis* culture supernatants and also during the protein purification scheme.

One Phadebas substrate tablet was suspended in 5 ml of distilled water. The substrate suspension was mixed thoroughly by vortexing and 1 ml was added to 10 µl of culture supernatant or α-amylase sample. The reaction mix was vortexed for 10 sec and incubated at 37°C for 15 min using a heating block with occasional mixing. The reaction was stopped by the addition of 200 µl of 0.5 M NaOH followed by vortexing for 10 sec. The assay was also carried out with 10 µl of distilled water for use as a reaction blank.

Solid material was pelleted by microcentrifugation (13 krpm, 10 *min*, room temperature) and the supernatants transferred to 1.5 ml cuvettes. The A_{620} for the sample/reaction blank was measured against a distilled water blank. The A_{620} of the reaction blank was subtracted from that of the sample and the activity determined using the standard curve provided with the kit. The activity value

obtained was multiplied by 4 as indicated in the Phadebas protocol for the urine assay system. One unit of α-amylase activity is defined as the amount of enzyme catalyzing the hydrolysis of 1 μmol of glycosidic linkage per minute at 37°C. For culture supernatants the specific activity of the enzyme was calculated by dividing the α-amylase activity by the OD_{660} of the culture at the time the sample was taken.

3.15 SDS polyacrylamide gel electrophoresis (SDS-PAGE) of proteins

Protein samples were analysed by denaturing SDS-PAGE (Mini-protean, Bio-Rad) according to the discontinuous method (Laemmli, 1970). The composition of the running and stacking gels are shown in Table 3.5. Protein samples (typically 10 μl) were mixed with an equal volume of x2 sample buffer (Appendix 3) and then boiled for 5-10 min immediately prior to loading. Electrophoresis was carried out using SDS-PAGE running buffer (Appendix 3) at 35 mA (constant current) per gel.

Table 3.5. Composition of gels used for denaturing SDS-PAGE

Reagent	Running gel (10% w/v)	Stacking gel (5% w/v)
Acrylamide:Bis (30% w/v ratio 37.5:1)	3.3 ml	1.0 ml
Tris.Cl (1.5 M, pH 8.8)	2.5 ml	-
Tris.Cl (0.15 M, pH 6.8)/0.1% SDS	-	5.0 ml
Distilled water	4.0 ml	-
SDS (10% w/v)	100 μl	-
TEMED	7 μl	3.5 μl
APS (10% w/v)	100 μl	50 μl

TEMED - N, N, N', N', - Tetramethylethylenediamine

APS - Ammonium persulphate

3.15.1 Gel staining

Proteins separated by SDS-PAGE were visualized by staining the gels with Coomassie Brilliant Blue. The gels were stained by submerging in Coomassie solution (Appendix 3) for 20-30 min with gentle agitation. Any dye not bound to protein within the gels was removed by washing several times with destain solution (Appendix 3), following which the gels were photographed and/or dried using a vacuum gel drier (Bio-Rad).

3.15.2 Fluorography

To enhance the signal from proteins labelled with ^{35}S methionine, the SDS-PAGE gels from pulse-chase experiments (3.28) were visualized by fluorography. Gels were fixed for 30 min in glacial acetic acid and then submerged in PPO solution (Appendix 3) for 30 min. The PPO in the gels was precipitated by washing the gels with distilled water and the gels were dried as before and exposed to X-ray film.

3.16 Detection of proteins using Western blotting

3.16.1 Transfer of proteins to cellulose nitrate membranes

Protein samples were separated by SDS-PAGE (3.15) and transferred (Trans-blotter, Bio-Rad) to a 0.4 μm cellulose nitrate membrane (Protan membranes, Schleicher and Schuell) using the protocol supplied by the manufacturer. Pre-stained size standards (Appendix 4) were included on the gels to allow the estimation of protein size and also to monitor the blotting efficiency. Where required, protein control samples (AmyL, AmyLQS50.5 or human serum albumin) were also included on the gels. Transfer was carried out at 70V (constant voltage) for 3 hours at 4°C in blotting buffer (Towbin *et al.*, 1979).

3.16.2 Detection of transferred proteins

Following transfer, the membrane was submerged in an excess of blocking buffer (Appendix 3) and incubated at room temperature for 30 min with gentle shaking. The blocking buffer was removed and replaced with 50 ml of fresh blocking buffer containing 50 μl of rabbit antiserum raised against AmyL (Anti-Termamyl, supplied by Novo Nordisk) or human serum albumin (Sigma), followed by incubation as before for 1 hour. The primary antibody solution was removed and the membrane was washed four times with wash solution I (Appendix 3), allowing 5 min at room temperature for each wash.

After the final wash, the membrane was submerged in 50 ml of wash buffer I containing 50 μl of porcine anti-rabbit IgG horse radish peroxidase conjugate (Dako Immunochemicals). The membrane was incubated in the secondary antibody for 1 hour and then washed four times with wash buffer II (Appendix 3). Finally, the membrane was rinsed with 10 mM Tris.Cl pH 7.5 and submerged in freshly prepared Western stain solution (Appendix 3) until protein bands became visible. The reaction was stopped by rinsing the membrane in distilled water and the stained blot was stored in the dark.

3.17 Isoelectric focusing (IEF) and zymography

The pI value can be used to determine the net charge of a protein; positively charged (basic) proteins have a high pI and negatively charged (acidic) proteins have a low pI. During IEF proteins maintain their native conformations and migrate through a pH gradient until they align themselves at the pH at which they have no net charge, *i.e* the pI. Carrier ampholytes are used to establish the pH gradient; proteins with a high pI move towards the cathode and those with a low pI move towards the anode.

3.17.1 IEF

α-Amylases present in culture supernatants were separated on the basis of charge using the Pharmacia Phast gel system with pre-cast IEF gels in the pH range 3-10. Sample volumes were typically 4-12 µl and the apparatus was operated according to the manufacturers recommendations.

3.17.2 Zymography

Following focusing, the gel was placed in a clean petri dish and overlayed with a mixture of 1% w/v molten agarose/1% w/v soluble starch which had been cooled to approximately 45°C. The agarose was allowed to solidify and the petri dish incubated at 37°C for 2-3 hours. Since proteins remain in their native conformations during IEF, the relative positions of the α-amylases within the gel were determined by staining the agarose/starch overlay with iodine vapour to reveal zones of starch hydrolysis.

3.18 Determination of the temperature optima of α-amylases

Cultures of *B. subtilis* were grown for 48 hours in 10 ml of 2xYT broth containing 6 µg/ml chloramphenicol and 1% w/v xylose. The cells were harvested by centrifugation (10000 *g*, 20 min, room temperature) and the supernatants removed and stored on ice. α-Amylase activity in diluted supernatant was determined at 37°C, 50°C, 65°C, 75°C, 90°C and 98°C using the Phadebas kit (3.14). Reaction blanks were also incubated at the different temperatures.

3.19 Determination of the pH optima of α-amylases

Culture supernatants (3.18) were diluted in 10 mM sodium phosphate buffer (Appendix 3) at the following pH; 5.7, 6.1, 6.7, 7.0, 7.3, 7.7 and 8.0. The α-amylase activities of the samples at the various pH values were determined (3.14) and expressed as a function of the actual pH, which was measured using a pH meter.

3.20 Determination of the thermostability of α-amylases

Culture supernatants (1 ml) from section 3.18 were incubated in a water bath at 90°C. At the following time intervals 10 µl samples were removed and stored on ice; 0, 2, 4, 6, 8, 10, 15, 20, 30 and 60 min. The α-amylase activities remaining in the samples were then determined (3.14) and represented the proportion of α-amylase which had not be subjected to irreversible thermoinactivation.

3.21 Determination of the stability of AmyLQS50 and AmyL

Cultures of *B. subtilis* DN1885 (pKS305) and DN1885 (pKS308) were grown for 48 hours in 2xYT (50 ml) containing 1% w/v xylose and 6 µg/ml chloramphenicol. The cells were removed by centrifugation and, to ensure the complete removal of cellular material, supernatants were then filter sterilized (0.45 µm Acrodisc-32 filters, Gelman Sciences).

Supernatants were incubated for 288 hours at 4°C in the presence or absence of 10 mM EDTA. At various time intervals, samples were removed and used for the determination of α-amylase activity (3.14) and Western blotting with anti-Termamyl antiserum (3.16).

After 194 hours at 4°C, an aliquot of each sample was removed and incubated at 37°C for 12 hours. Following incubation at 37°C, the samples were assayed for α-amylase activity and subjected to Western blotting as before.

3.22 Isolation of chromosomal DNA from *B. subtilis*

Chromosomal DNA was isolated from *B. subtilis* strains using the IGi Genomic DNA kit (Immunogen International). The cells from 300 µl of an overnight culture of *B. subtilis* were harvested by microcentrifugation (13 krpm, 5 min, room temperature), the supernatant removed and the cells resuspended in 300 µl of solution I. Three hundred microlitres of solution II were then added, the solutions mixed by inversion and incubated at 55°C for 30 min. After cooling to room temperature, 150 µl of solution III was added, the mixture vortexed and then incubated at room temperature for 10 min.

Solid material was pelleted by microcentrifugation (13 krpm, 10 min, room temperature) and the supernatant was transferred to a fresh tube. The chromosomal DNA was precipitated (3.3), resuspended in 50 µl of distilled water and analyzed by agarose electrophoresis (3.6).

3.23 Southern hybridization of *B. subtilis* chromosomal DNA

To verify the integration of plasmids encoding chimeric α-amylases into the chromosome of *B. subtilis* DN1885, chromosomal DNA isolated from various strains was subjected to Southern hybridization using the *B. subtilis xylR* gene as a probe.

3.23.1 Gel preparation

Isolated chromosomal DNA (3.22) was digested with *Bgl*II (3.5) and electrophoresed on a 1% w/v agarose gel (3.6). The DNA fragments were depurinated by soaking the gel in 0.25 M HCl for 25 min. The gel was then rinsed in distilled water and submerged in denaturation solution (Appendix 3) for 30 min. Following denaturation, the gel was rinsed with distilled water and placed in neutralizing solution twice (Appendix 3) for 15 min with gentle shaking.

3.23.2 Transfer of DNA to a nylon membrane

The DNA was transferred from the gel to a nylon membrane (Hybond-N, Amersham International plc) using the capillary transfer method (Southern, 1975) with x20 SSC (Appendix 3) as the transfer buffer. The membrane was rinsed in x2 SSC, air dried and the DNA fixed by baking at 80°C for 2 hours.

3.23.3 Hybridization

Hybridization was carried out in a hybridization oven (Hybaid). The membrane was placed inside a glass tube, 25 ml of pre-hybridization solution (Appendix 3) added and incubated for 2 hours at 65°C. The pre-hybridization solution was removed and replaced with 25 ml of hybridization solution (Appendix 3) containing the denatured ^{32}-P labelled *xylR* probe (3.24). Hybridization was carried out overnight at 65°C and, after removal of the probe solution, the membrane was washed twice at 65°C for 30 min using x3 SSC/0.1% SDS. The membrane was wrapped in Saran wrap and exposed to X-ray film.

3.24 DNA probe labelling

To make the probe for Southern hybridization, plasmid pCJ92 (10 μg) was digested with *Sph*I (3.5) and the ~1.6 kbp fragment encoding the *xylR* gene was gel purified (3.7). The purified *xylR* fragment was labelled with α-^{32}P dCTP using the Rediprime DNA labelling kit (Amersham International plc) according to the manufacturers instructions. This kit uses random sequence hexanucleotides

to prime DNA synthesis, resulting in a probe with an extremely high specific activity (>10^9 cpm/μg of DNA). Following labelling, the probe was stored at -20°C and was stable for up to 1 month (approximately 2 half-lives).

3.25 Autoradiography and X-ray film development

Dried gels and membranes were exposed to X-ray film for various lengths of time depending on the isotope used. Labelled proteins (^{35}S) were exposed to X-ray film at -20°C whereas, labelled DNA (^{32}P) was exposed at room temperature. Following exposure, the autoradiographs were developed using Photosol developing and fixing solutions (GRI), according to the manufacture's instructions.

3.26 Purification of α-amylases from culture supernatants

3.26.1 Growth

Colonies of *B. subtilis* from a fresh penassay agar plate were used to inoculate 500 ml of 2xYT broth (containing 6 μg/ml chloramphenicol and 1% w/v xylose) in duplicate. The cultures were grown in 3 l conical flasks to facilitate good aeration. The cultures were grown for 48 hours at 37°C with shaking at 100 rpm and the cells were harvested by centrifugation (10000 *g*, 30 min, 4°C).

3.26.2 Ammonium sulphate precipitation of proteins

The duplicate culture supernatants from each strain were combined and proteins precipitated by the step-wise addition of ammonium sulphate with continuous stirring. In the initial precipitation step, proteins were precipitated by the addition of ammonium sulphate to 50% saturation (313 g per litre of supernatant) and pelleted by centrifugation (10000 *g*, 1 hour, 4°C). The supernatant was removed and the pelleted protein resuspended in 9 ml of 5 mM sodium phosphate buffer pH 7.2 (Appendix 3).

Additional ammonium sulphate (214g per litre of supernatant) was added to the supernatants to increase the concentration to 80% saturation and again the precipitated protein was pelleted by centrifugation. The supernatant from the 80% precipitation was discarded and the protein resuspended in 12 ml of sodium phosphate buffer. The precipitated proteins were analysed by SDS-PAGE (3.15).

3.26.3 Dialysis

Before use, the dialysis tubing (2.4 nm pore, BDH) was prepared by boiling for 10 min in a solution of 2% w/v sodium bicarbonate/1 mM EDTA (pH

8.0). Once cool, the tubing was boiled in 1 mM EDTA (pH 8.0) for 10 min and stored at 4°C in the same solution containing 0.05% sodium azide.

To remove residual ammonium sulphate and any salts carried over from the growth medium, protein solutions were dialysed against a large volume (~1 l) of 5 mM sodium phosphate buffer pH 7.2 for 48 hours, with several buffer changes. All the dialysis steps were carried out at 4°C to preserve enzymatic activity.

3.26.4 Separation of proteins using the Rotofor cell (Bio-Rad)

The Rotofor cell is a preparative IEF cell which separates proteins on the basis on charge. The dialyzed protein solutions were mixed with ampholytes in sterile distilled water to a final volume of 55 ml. The ampholytes (Biolyte 3/10, Bio-Rad) were added to a final concentration of 2% w/v and provided a pH range of 3 to 10.

The Rotofor apparatus was assembled and operated in accordance with the protocol provided by the manufacturer. The protein/ampholyte solution was added to the cell and focusing carried out for 4-6 hours. Focused proteins were harvested in 20 fractions each with a volume of approximately 2.75 ml. The lower fraction numbers were closest to the anode (pH 3) and, as a consequence, contained proteins with an overall negative charge. Conversely, positively charged proteins moved towards the cathode (pH 10) and were harvested with the higher fractions. Fractions were assayed for α-amylase activity (3.14) and analyzed by SDS-PAGE (3.15).

3.26.5 Removal of contaminating ampholytes and sample concentration

The Rotofor fractions containing the majority of the α-amylase activity were dialyzed against large volumes of 1 M NaCl to remove the electrostatically bound ampholytes by ion exchange. Dialysis was carried out for 72 hours in dialysis cassettes (Slide-a-Lyzer cassettes, Pierce) which had a 10 kDa molecular weight cut-off point. Following the removal of ampholytes, the samples were dialysed for 48 hours against 5 mM sodium phosphate buffer pH 7.2 to remove NaCl.

The α-amylase samples were then concentrated by dialysis against 20% w/v polyethylene glycol 6000 in 5 mM sodium phosphate buffer pH 7.2 for 1-2 hours. Dialysis against polyethylene glycol resulted in the movement of water molecules by osmosis from the protein sample into the polyethylene glycol solution with the consequential increase in protein concentration. Once dialysis against polyethylene glycol was complete the volume of the protein samples was approximately 200 μl.

3.27 Determination of protein concentration

Following purification or prior to *in vitro* wall binding studies, the concentration of proteins in α-amylase preparations was determined using the bicinhoninic acid (BCA) protein assay kit (Pierce). The assay was carried out according to the manufacturer's instructions using the "standard" protocol.

3.28 Pulse-chase and immunoprecipitation

The secretion kinetics of α-amylases were investigated using pulse-chase experiments. Cultures were grown in modified Spizizens minimal medium (SMM, Appendix 3) with 1% w/v ribose as the carbon source and 1% w/v xylose to induce the synthesis of α-amylase. Ribose was used as the carbon source because, whilst permitting relatively high growth rates, this sugar did not cause a significant degree of catabolite repression of α-amylase expression. Hence, the combination of ribose and xylose permitted the synthesis of α-amylase by exponentially growing cultures.

3.28.1 Growth

Single colonies of each strain from a penassay agar plate containing chloramphenicol were used to inoculate 18 ml of SMM which had been pre-warmed to 37°C. The cultures were incubated at 37°C with shaking at 180 rpm and growth was monitored by measuring the OD_{660}. When the OD_{660} reached 0.7 to 0.9, a 15 ml sample was removed and transferred to a sterile plastic universal tube which had been pre-warmed to 37°C.

3.28.2 Pulse-chase

Culture samples were incubated in a water bath at 37°C and proteins were labelled by the addition of 100 μCi of L-^{35}S methionine (Amersham International plc). After 1 min, a 600 μl "chase" of unlabelled methionine (25 mg/ml) was added to stop the incorporation of isotope. Immediately after the chase (0 min), and at time intervals thereafter (0.5, 1.0, 2.0, 3.0, 4.0, 5.0, and 6.0 min), two 600 μl samples were removed from each culture.

The total protein (cell associated + supernatant) from one 600 μl sample was precipitated by adding directly to 600 μl of 20% w/ v trichloroacetic acid (TCA) on ice. The second sample was filtered (0.45 μm PVDF filters, Whatman) into 20% w/v TCA to precipitate released proteins present in culture supernatants. The samples were left on ice for 1 hour and then microcentrifuged (13 krpm, 15 min, room temperature) to pellet solid material. The pellets were washed with 1 ml of acetone to remove residual TCA and dried under vacuum.

The samples were resuspended in 100 µl of lysis buffer (Appendix 3) and incubated at 37°C for 5 min. Following incubation, 10 µl of 10% w/v SDS was added to each sample and proteins solubilized by boiling for 10 min. After cooling to room temperature, 900 µl of x1 STD (Appendix 3) was added and the samples were incubated on ice for 30 min. Solid material was pelleted by microcentrifugation (13 krpm, 5 min, room temperature) and the supernatants transferred to new tubes.

3.28.3 Immunoprecipitation

Anti-Termamyl antiserum (10 µl) was added to the supernatants and the tubes were incubated at 4°C overnight to allow immunoprecipitation. Solid material was pelleted by centrifugation and the supernatants, containing immunoprecipitated complexes, were transferred to tubes containing 5 mg of Protein A Sepharose beads (Pharmacia) in 100 µl of x1 STD. The tubes were incubated at room temperature for at least 1 hour with shaking to allow the binding of antibody-antigen complexes to the Protein A Sepharose beads.

The Sepharose beads were pelleted by microcentrifugation (13 krpm, 1 min, room temperature) and washed 4 times with x1 STD. After the final wash the supernatants were removed and beads resuspended in 40 µl of x2 sample buffer (Appendix 3). The samples were boiled for 5 min to release the immunprecipitated α-amylases from the Sepharose beads and analyzed by SDS-PAGE (3.15) with subsequent fluorography (3.15.2). The relative amounts of precursor and mature α-amylases were determined by analyzing the resultant autoradiographs using the Gel Documentation system and Molecular Analyst software (Bio-Rad).

3.29 Measurement of the unfolding/refolding transition of α-amylases

The degree of resistance to proteolytic digestion can be used to determine the folded state of proteins such as AmyL. When folded into its native conformation, AmyL is highly resistant to the action of proteases found in culture supernatants. However, once partially or completely unfolded, AmyL becomes sensitive to proteolytic degradation. Protease sensitivity was used to monitor the unfolding/refolding transition of Amy-LQS50.5 and AmyL. The experimental details are described below. The procedure was based on the method described previously for *B. subtilis* levansucrase (Chambert *et al.*, 1995).

3.29.1 Unfolding

The α-amylases were unfolded (denatured) by mixing 1.5 µl of purified AmyLQS50.5 (515 ng/µl) or AmyL (685 ng/µl) with 38.5 µl of a solution of 6 M

guanidine hydrochloride (pH 7.0) containing 1 mM EDTA. The presence of EDTA at this stage was essential for the unfolding of α-amylases due to the role of Ca^{2+} in structural integrity. The unfolding mixtures were incubated at 47°C and at five different time intervals 4 µl samples were removed and added to 1.5 ml tubes containing 96 µl of acetate buffer (0.1 M sodium acetate at pH 7.0 containing 0.2 mM calcium chloride) and 20 µl of a 1 mg/ml solution of subtilisin Carlsberg (Sigma). The reaction mixes were incubated at 25°C for 5 min to allow the proteolytic degradation of unfolded α-amylase and, after the incubation period, the protease was inactivated by the addition of 5 µl of PMSF (0.1 M solution in ethanol).

3.29.2 Refolding

To initiate refolding, 1 ml of acetate buffer (0.1 M sodium acetate at pH 7.0 containing 0.2 mM calcium chloride) was added to the remainder of the unfolding mixtures (20 µl) to dilute the denaturant from 6 M to 120 mM. The mixtures were incubated at 25°C and at different time intervals 100 µl samples were removed and added to 20 µl of subtilisin solution. Proteolytic digestion of unfolded protein was allowed to proceed for 5 min at 25°C after which the subtilisin was inactivated by the addition of 5 µl of PMSF, as before.

3.29.3 Determination of α-amylase activity

Since the folded α-amylases were resistant to degradation by subtilisin, the α-amylase activity of a particular sample represented the amount of fully folded protein present. Therefore, the α-amylase activity of the different samples was determined using the Amylase-3 substrate (Sigma). To each of the unfolding/refolding samples (125 µl) from above, 575 µl of acetate buffer was added to make a final volume of 1 ml.

The lyophilized Amylase-3 substrate was reconstituted in 3.5 ml of distilled water and 300 µl of this substrate solution was then added to each sample. The tubes were vortexed briefly to mix the reagents and then incubated at 37°C for 5 to 20 min. A note was made of the exact incubation time for each sample so that the rate of change of absorbance could be calculated. Following incubation at 37°C, the A_{405} was measured against a distilled water blank using an LKB Ultrospec II spectrophotometer. The rate of change of absorbance was determined and used to calculate the units of α-amylase per ml of solution, as described in the protocol provided by the manufacturer, taking into account the relevant dilution factors. The units of α-amylase per ml against time for the different enzymes were then plotted to allow the unfolding/refolding transitions of the enzymes to be compared.

3.30 The *in vitro* cell wall binding (IVWB) assay

This assay was developed to assess the ability of proteins in their native conformations to bind to *B. subtilis* cell walls *in vitro*. The cell walls used were isolated from *B. subtilis* 168 grown in phosphate replete conditions as described previously (Merad *et al.*, 1989) and had teichoic and teichuronic acid compositions of 0.927 and 0.145 μmole/mg of wall, respectively. Teichoic acids may be substituted by D-alanyl esters which have the effect of neutralizing the negative charge mediated by the phosphate groups (2.10). Consequently, before the IVWB assay was carried out the walls were treated to lower the degree of alanylation of teichoic acids and therefore, increase the overall negativity of the wall. Wall samples (26 mg) were washed several times in 500 μl aliquots of Tris.HCl pH 8.0 (Appendix 3). The washing was carried out at pH 8.0 because the linkage between the alanyl ester and the C-2 position of the glycerol residue of teichoic acid is labile, even at relatively mild alkaline pH (Shabarova *et al.*, 1962). Washing was also carried out to remove any non-covalently bound molecules or ions from the wall preparation. The initial concentration of Tris.HCl used was 1 M and this was reduced in 10-fold steps in subsequent washes, until the final concentration was 0.1 mM. The wall was then washed three times with 500 μl of distilled water and resuspended in 200 μl of distilled water.

The IVWB assay was carried out using purified AmyLQS50.5 and AmyL. A pure preparation of human serum albumin (HSA) was obtained from Sigma and prior to the assays, the exact concentrations of all protein samples were determined using the BCA protein assay kit (3.27). The protein samples were diluted to 25 ng/μl in 5 mM sodium phosphate buffer pH 7.2 (Appendix 3) and the IVWB assay was carried out in individual 1.5 ml microcentrifuge tubes containing the reaction components indicated in Table 3.6.

Sample I was the wall binding reaction with both wall and protein present. Sample II represented the total amount protein put into the system and did not include any wall material. Finally, sample III was used as a negative control to confirm that the proteins detected were specific and not caused by cross-reacting material in the wall preparation. After all the components of the reactions had been mixed together the tubes were incubated at 37°C for 15 min to allow the proteins and the cell walls to interact.

Continued

Table 3.6. Components of the IVWB assay

	Sample I	Sample II	Sample III
Protein solution	50 µl	50 µl	-
Wall suspension	35 µl	-	35 µl
Distilled water	-	35 µl	50 µl

Following incubation, the cell walls in samples I and III were pelleted by microcentrifugation (13 krpm, 5 min, room temperature) and the supernatants transferred to fresh tubes. Aliquots (60 µl) of samples I and II were then added to equivalent volumes of x2 sample buffer (Appendix 3) in separate tubes. The pelleted walls, termed samples IW and IIIW, respectively, were washed three times with 200 µl of distilled water and then resuspended in distilled water to a final volume of 60 µl. Double-strength sample buffer (60 µl) was then added to the IW and IIIW wall suspensions which were subsequently boiled for 10 min to release non-covalently bound proteins from the wall. After incubation at 37°C for 15 min 60 µl of sample II was also removed and added to an equal volume of x2 sample buffer.

Samples (20 µl) were subjected to SDS-PAGE (3.15) and Western blotting (3.16) and prior to gel loading the samples containing wall material (IW and IIIW) were microcentrifuged briefly to pellet solid material. Pre-stained molecular size standards and protein controls were also included on the gels. The protein controls consisted of 10 µl of Termamyl (25 ng/µl) or HSA (25 ng/µl) mixed with an equal volume of x2 sample buffer. The primary antiserum used for the Western blots was anti-Termamyl in the case of AmyLQS50.5 and AmyL, and anti-HSA (Sigma) for HSA.

CHAPTER 4

THE CONSTRUCTION OF PLASMID VECTORS FOR THE STUDY OF PROTEIN SECRETION IN *BACILLUS SUBTILIS*

4.1 Introduction

During this study vector systems were developed to facilitate the study of post-translocational secretion events. Plasmid pSX63, provided by Novo Nordisk, was based on the *Staphylococcus aureus* plasmid pUB110. A map of pSX63 is shown in Appendix 5 and this plasmid allowed the inducible expression of AmyL in response to the presence of xylose in the growth medium (2.19). The large *Eco*RI to *Bgl*II DNA fragment of pSX63 encodes all the elements required for the inducible expression of α-amylase genes and was used as a cassette for the inducible expression of various α-amylase genes. The smaller *Eco*RI to *Bgl*II fragment of pSX63 contains the elements necessary for the replication of this plasmid in *B. subtilis*.

4.2 Engineering pSX63 to remove an unwanted *Pst*I restriction site

The chimeric α-amylase genes used to study the later stages in the export process were constructed as *Pst*I to *Hind*III fragments to make use of the signal sequence and transcription terminator of the *amyL* gene. pSX63 contains two *Pst*I and one *Hind*III restriction sites (Appendix 5). The unique *Hind*III site is located at the end of the *amyL* gene, downstream from the transcription terminator. One of the *Pst*I sites is located in the *amyL* signal sequence and this restriction site, together with the *Hind*III site, were to be used to clone the constructed α-amylase genes. The other *Pst*I restriction site is located between the *cat* and *xylR* genes and presented a problem because, as long as this site was present, α-amylase genes could not be cloned into the *Pst*I site located within the signal sequence of the *amyL* gene. Therefore, steps were taken to remove this restriction site from pSX63.

The unwanted *Pst*I restriction site was removed from pSX63 in three stages. As indicated in the plasmid map (Appendix 5), pSX63 contains two *Cla*I restriction sites, one upstream from the promoter/operator and the other within the region coding for the mature AmyL protein. Removal of the small *Cla*I fragment from pSX63 (approximately 700 bp) would remove the inducible promoter, the DNA encoding the signal peptide and consequently the *Pst*I site to be used to clone the α-amylase genes.

In the first stage pSX63 was restricted with *Cla*I (3.5) and the large *Cla*I DNA fragment purified from an agarose gel (3.7). This fragment was religated (3.8) at the now unique *Cla*I site and used to transform naturally competent *B. subtilis* DN1885 (3.9). The resultant plasmid, pKS200 (Appendix 5), contained a truncated and non-functional *amyL* gene with a unique *Pst*I restriction site located between the *cat* and *xylR* genes.

The second and third stages were carried out by C L Jensen at Novo Nordisk and involved the removal of the unique *PstI* restriction site of pKS200 and the subsequent replacement of the truncated *amyL* gene with a functional version. pKS200 was linearized with *Pst*I and the overhanging ends of the *Pst*I restriction site were trimmed back to form blunt ends using S1 nuclease. The blunt ends produced were then ligated and the DNA used to transform *B. subtilis* DN1885 as before. In the third and final stage the *Eco*RI to *Hind*III *amyL* fragment from pSX63 was used to replace the truncated *amyL* fragment of pKS200Δ*Pst*I. The resultant plasmid, pCJ92 (Appendix 5), possessed a unique *Pst*I restriction site within the *amyL* signal sequence and synthesized α-amylase only in the presence of xylose. DNA sequencing of the region between the *cat* and *xylR* genes was used to verify that only the *Pst*I restriction site was deleted.

4.3 Construction of an *E. coli*/*B. subtilis* shuttle vector

Under most transformation conditions the uptake of plasmid DNA in *B. subtilis* is not normally efficient enough to permit cloning directly in this host. The low transformability of *B. subtilis* is related to a requirement for the presence of plasmid multimers in the transforming DNA preparation (Canosi *et al.*, 1978). Therefore, alternative strategies to introduce recombinant plasmids into *B. subtilis* had to be found. One such approach is to use a shuttle vector which is capable of replicating in *B. subtilis* and also in a more easily transformed host, such as *E. coli*. Therefore, a cloning strategy was devised to construct a shuttle vector containing the inducible secretion cassette which would allow the cloning of the chimeric α-amylase genes in *E. coli* and, once amplified, the recombinant plasmids containing the chimeric genes could be transferred to *B. subtilis*.

The *E. coli* replicon and antibiotic resistance gene chosen for use in the shuttle vector was that from plasmid pUC19 (Yanisch-Perron *et al.*, 1985). When in *E. coli*, this plasmid has a copy number of approximately 500 per chromosome and confers resistance to ampicillin due to the presence of the gene encoding β-lactamase. A shuttle vector was constructed by cloning the origin of replication and β-lactamase gene from pUC19 into a suitable location on pCJ92. This was achieved with PCR by amplifying over the *Ssp*I (2501 bp) and *Sap*I (683 bp) restriction sites of pUC19 using primers KS-3 and KS-4 (Appendix 2), respectively. These primers were designed with 5' extensions to introduce *Bgl*II restriction sites into the ends of the resulting PCR product to facilitate cloning into the unique *Bgl*II site of pCJ92, thus generating a plasmid able to replicate in *E. coli* and *B. subtilis*.

PCR was carried out (3.12) and the size of the resultant PCR product, 1.85 kbp, was verified by agarose electrophoresis (3.6) (Fig. 4.1). The PCR product

was gel purified and (3.7) and digested with *Bgl*II (3.5). pCJ92 DNA was also digested with *Bgl*II and both fragments were gel purified. DNA fragments were concentrated by precipitation (3.3), ligated (3.8) and used to transform chemically competent *E. coli* XL1-Blue (3.10.4), selecting on penassay agar plates containing ampicillin. Plasmid DNA from one putative clone was isolated using the STET buffer method (3.4.1) and analysed by double restriction digestion with *Eco*RI/*Hind*III and *Pst*I/*Hind*III with subsequent agarose electrophoresis (Fig. 4.2). pSX63 was also analysed in parallel.

Digestion of pSX63 with *Eco*RI and *Hind*III produced two fragments, the smaller of which was 2.2 kbp in size and corresponded to the inducible promoter, signal sequence and mature region of the *amyL* gene. A fragment of the same size was also liberated from the putative clone. Additionally, the plasmids were digested with *Pst*I and *Hind*III to liberate, amongst others, the 1.7 kbp DNA fragment corresponding to the mature region of *amyL*. Therefore, the restriction analysis carried out indicated that the cloning had proceeded as planned and the shuttle vector constructed, termed pKS301, was approximately 8.8 kbp in size (Appendix 5). This clone was selected for the large scale preparation of plasmid DNA (3.4.3).

The production of α-amylase by *E. coli* XL1-Blue (pKS301) was investigated by streaking colonies onto penassay agar plates containing ampicillin and 1% w/v soluble starch, in the presence or absence of 1% w/v xylose. Following incubation overnight at 37°C, the production of a functional α-amylase by the colonies was revealed by staining with iodine vapour. When xylose was included in the plates, large zones of starch hydrolysis were observed around colonies containing pKS301 indicating the presence of the inducible α-amylase gene. However, there was also some evidence of starch hydrolysis around equivalent colonies in the absence xylose suggesting that the promoter was not fully repressed in this host. Starch hydrolysis was not observed around colonies of *E. coli* XL1-Blue without pKS301. Since the zones of starch hydrolysis were much larger in the presence of xylose than in its absence, the production of α-amylase under non-inducing conditions in *E. coli* XL1-Blue was not investigated further.

In summary, pKS301 was constructed from the parent plasmid pCJ92 to allow the relatively high level expression of various α-amylase genes in *B. subtilis*. The shuttle plasmid contains the origin of replication from pUB110 together with an amplified copy of the origin of replication from pUC19. pKS301 also encodes a β-lactamase gene to allow selection in *E. coli* and a *cat* gene for selection in *B. subtilis*. *E. coli* XL1-Blue (pKS301) was able to produce relatively large quantities of α-amylase in the presence of xylose which could be detected

in the culture medium by a simple plate assay. However, whether this enzyme was secreted efficiently across the outer membrane or appeared in the medium as a consequence of leakage from the periplasm of *E. coli* was not investigated.

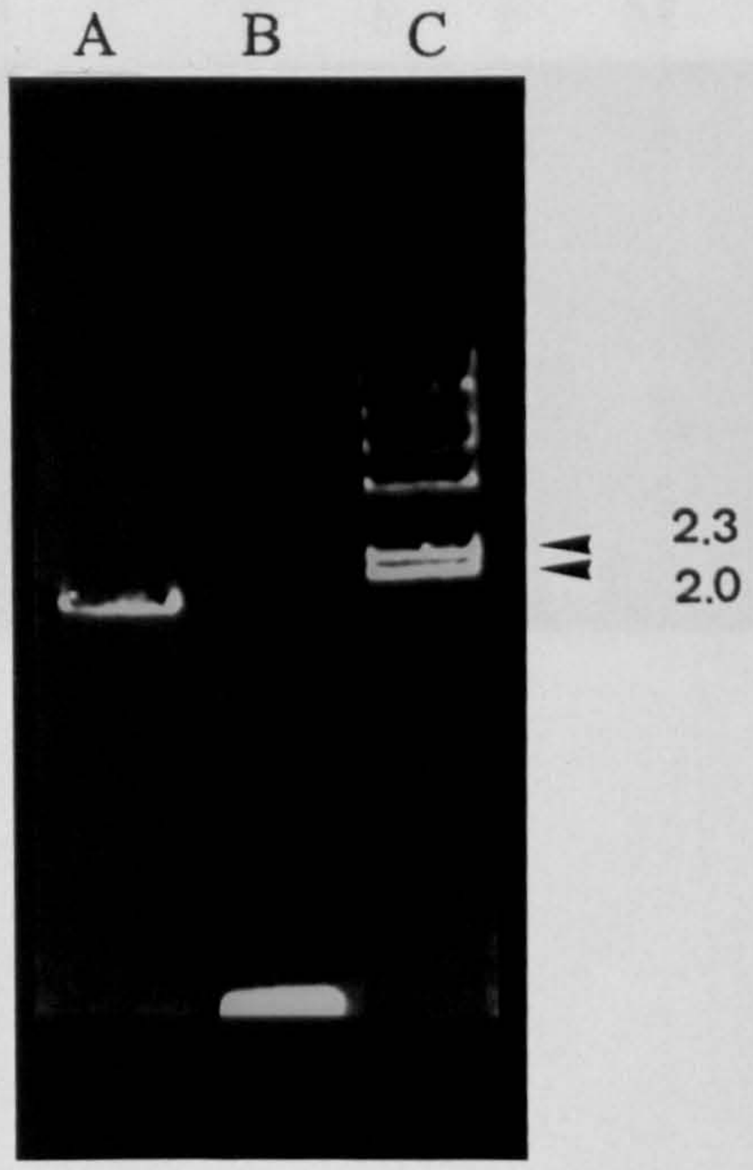


Fig. 4.1. Agarose electrophoresis of the pUC19 KS3/4 PCR product and control reaction.

Key:

A - PCR product

B - Negative PCR control (no added template)

M - λ *Hind*III markers (kbp)

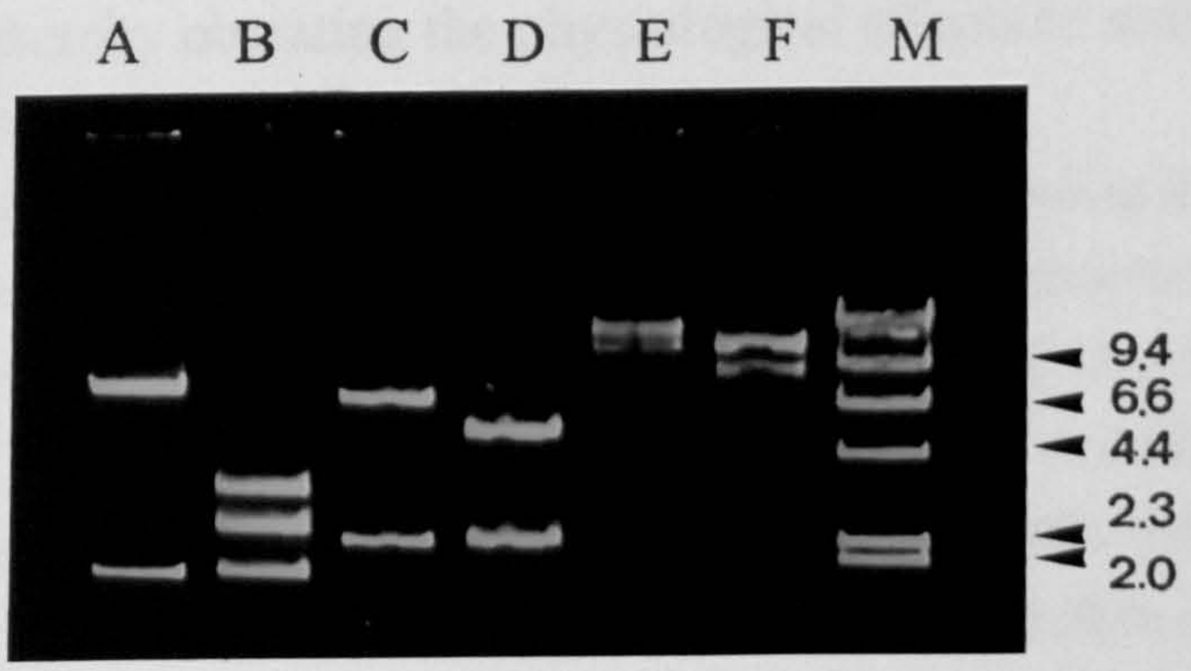


Fig. 4.2. Restriction analysis of pKS301 and pSX63.
Key:

A - pKS301 *Pst*I / *Hind*III digest
B - pSX63 *Pst*I / *Hind*III digest
C - pKS301 *Eco*RI / *Hind*III digest
D - pSX63 *Eco*RI / *Hind*III digest
E - Uncut pKS301
F - Uncut pSX63
M - λ *Hind*III markers (kbp)

4.4 Construction of a *B. subtilis* integration vector

Certain plasmids which are capable of replicating in *B. subtilis*, such as those based on pUB110, are known to be structurally and segregationally unstable in this host due, at least in part, to the absence of a minus origin (Gryczan, 1982; Bron and Luxen, 1985). To avoid problems associated with plasmid instability, a vector was constructed to facilitate the integration of xylose-inducible α-amylase cassettes into the chromosome of *B. subtilis*. Additionally, this allows the maintenance of the α-amylase genes in the absence of selection, thereby obviating the physiological affects of antibiotics such as chloramphenicol on the cell wall.

The integrational plasmid was constructed by cloning a DNA fragment from pCJ92 encoding the xylose-inducible α-amylase cassette system into pBluescript SK⁻ (Appendix 5). pBluescript SK⁻ is a 2.96 kbp *E. coli* plasmid based on pUC19 which confers resistance to ampicillin and allows blue/white selection of recombinants (3.11). The fragment cloned into this *E. coli* vector contained all the elements necessary for the xylose-inducible expression of α-amylase but not an origin of replication which was functional in *B. subtilis*.

pKS301 DNA was digested with *Eco*RI and *Bgl*II and the large fragment (approximately 6 kbp), encoding the xylose-inducible α-amylase cassette, was gel purified. This fragment was ligated with *Eco*RI and *Bam*HI digested pBluescript SK⁻, the latter enzyme producing ends compatible with *Bgl*II. The ligation mixture was used to transform competent *E. coli* XL1-Blue (3.10.4) and transformants were selected on penassay agar containing ampicillin, X-Gal and IPTG. After incubation overnight at 37°C, white (recombinant) colonies were selected for the small scale extraction of plasmid DNA using the STET method (3.4.1). The presence of the correct fragments within the isolated plasmids was verified by restriction analysis; more than 60% of the plasmids tested contained an insert of the expected size. One clone was selected for the large scale purification of plasmid DNA (3.4.3) and this plasmid was called pKS401 (Appendix 5).

In conclusion, pKS401 is approximately 9 kbp in size and encodes the xylose-inducible α-amylase secretion cassette. This plasmid integrates into the chromosome of *B. subtilis* by homologous recombination between the plasmid- and chromosomally-located *xylR* genes (7.4) and was constructed with the aim of increasing the overall stability of the expression system.

CHAPTER 5

THE CONSTRUCTION OF GENES ENCODING CHIMERIC α-AMYLASES WITH ALTERED NET CHARGE

5.1 Introduction

Following translocation across the cytoplasmic membrane, *B. subtilis* exoproteins which function in the extracellular milieu must traverse the cell wall in what is the last stage of the export pathway before becoming truly extracellular. Other exoproteins, such as WAP's, appear to exert their function within the wall cylinder and therefore, must remain localized at this site. As an exoprotein enters the cell wall cylinder, the degree to which it interacts with the components of the wall is determined by the structure and properties of the exoprotein and also the composition of the wall (2.13). Hence, certain *B. subtilis* exoproteins have evolved to reduce interactions with the cell wall and promote unhindered egress from the wall cylinder, whilst others, such as WAP's, have developed features which ensure that they remain closely associated with the wall (2.11).

The cross-linked peptidoglycan-anionic polymer complex of the cell wall has physico-chemical properties which promote interactions between the wall and secretory proteins. The rigid glycan strands and cross-linking peptides constitute a mesh-like matrix with considerable sieving properties (2.13.1) which may retard the passage of larger proteins more than smaller ones. Further, the covalently attached anionic polymers impart a negative charge and consequently confer ion exchanger properties to the wall (2.13.2) which may affect the passage of charged proteins through the wall cylinder.

Whilst moving through the matrix of the wall, charged regions of exoproteins can be expected to interact strongly with the immobilized anionic groups of the wall. On a simple charge-charge basis, basic proteins would be expected to bind strongly to the attached anionic polymers and therefore, to be retained within the wall cylinder. Conversely, electrostatic repulsion between acidic exoproteins and the negatively charged groups of the wall may either promote the expulsion of the exoprotein from the outer surface of the wall cylinder or, alternatively, prevent it from entering the cylinder at the inner surface.

A number of lines of evidence suggest that exoprotein charge is an important factor in determining the degree to which exoproteins interact with, and bind to, the *B. subtilis* cell wall during export. For example, the majority of rapidly exported proteins from *B. subtilis* W23 have pI's in the range 6.5 to 7.0 (Coxon, 1990) which may represent the tailoring of the properties of exoproteins to reduce charge interactions with the cell wall (2.13.2). Conversely, the *B. subtilis* N-acetylmuramoyl-L-alanine amidase and modifier protein, which bind to cell wall by an interaction with teichoic acids have, in addition to N-terminal wall binding motifs, pI values greater than 10.0 (Lazarevic *et al.*, 1992).

Similarly, the Epr extracellular protease of *B. subtilis* is putatively wall bound (Pero and Sloma, 1993) and has, in its C-terminus, potential wall binding domains and also an unusually large number of positively charged amino acids which may facilitate the localization of this protease in the wall cylinder.

To investigate the affect of charge on the secretion of exoproteins, variants of a well secreted protein were constructed. The model protein chosen for this purpose was the α-amylase from *B. licheniformis*, AmyL, which was selected because this enzyme is known to be expressed and secreted normally in *B. subtilis* (Ortlepp *et al.*, 1983; Laoide *et al.*, 1989), and is closely related to the α-amylases of *B. amyloliquefaciens* (AmyQ) and *B. stearothermophilus* (AmyS) (Table 5.1).

The relatedness of AmyL, AmyQ and AmyS at the amino acid (and nucleotide) level allows amino acids from one enzyme to be replaced with the corresponding amino acids from the other enzymes to produce chimeric α-amylases which retain catalytic activity. This has been demonstrated by several authors who have reported the construction of *in vivo* and *in vitro* derived chimeras of these α-amylases. A number of chimeric α-amylase genes encoding active enzymes have been constructed by *in vivo* genetic engineering (Gray *et al.*, 1986; Diderichsen *et al.*, 1987; Jørgensen *et al.*, 1990). This technique is dependent upon homologies between different α-amylase genes which, when encoded on the same plasmid and introduced into a suitable host, recombine at homologous regions to produce chimeric genes. *In vivo* recombination has the advantage that it allows the direct selection of recombination events leading to the secretion of catalytically active chimeric α-amylases.

Another approach for the construction of chimeric α-amylase genes employs more standard *in vitro* molecular cloning techniques to restrict and ligate regions of genes at specific and predefined restriction sites (Suzuki *et al.*, 1989). However, this technique is limited by the requirement for suitable restriction sites within the α-amylase genes. What is clear is that, in most cases, the properties of the products of the chimeric α-amylase genes, generated by both *in vivo* and *in vitro* techniques, differed when compared to the relevant wild type α-amylases (Gray *et al.*, 1986; Diderichsen *et al.*, 1987; Suzuki *et al.*, 1989). Therefore, to generate charge variants of AmyL, it was decided to use the genes encoding the related α-amylases to construct chimeric α-amylase genes encoding enzymes with predictable changes in overall charge.

Table 5.1. Characteristics of the α-amylases from *B. licheniformis*, *B. amyloliquefaciens* and *B. stearothermophilus*

Organism	Gene	Size (kbp)	Mature protein amino acids	Mature protein size (kDa)	% Similarity
B. licheniformis	*amyL*	1.54	483	55.1	-
B. amyloliquefaciens	*amyQ*	1.55	484	54.9	79.3
B. stearothermophilus	*amyS*	1.65	515	58.8	63.4

The sizes of the respective genes in Table 5.1 are measured from the start codon to the stop codon since the chimeric α-amylases were expressed using a promoter system other than the native promoters. The percentage similarity of AmyQ and AmyS to AmyL were calculated using Megalign, part of the Lasergene biocomputing software (DNASTAR Inc, Wisconsin, USA).

5.2 The design of chimeric α-amylases

To utilize the well developed knowledge of enzyme engineering at Novo Nordisk, and in particular that relating to the engineering of α-amylases, the sequences of the chimeric α-amylases were designed by C L Jensen in the Bacterial Gene Technology department of Novo Nordisk. A range of chimeric enzymes were designed *in silico* using PC/GENE (IntelliGenetics Inc., USA). Various regions of the mature AmyL protein were replaced with corresponding regions of AmyQ and AmyS. These regions were identified on the basis of charged amino acid residues and were replaced with equivalent amino acids from the other α-amylases to either increase or decrease the calculated pI of the resultant chimeras when compared to wild type AmyL. The pI values for the various proteins were calculated using the titration function of the Protean protein analysis package which was also part of the Lasergene biocomputing software.

By swapping regions of the α-amylase genes encoding charged amino acids, a range of theoretical chimeric enzymes were generated. As indicated in Fig. 5.1, the most extreme (hypothetical) α-amylase charge variants which could be generated were termed AmyLQS45 and AmyLQS46, with overall negative and positive charges, respectively. Using this method the calculated pI values of the extreme cases could be reduced to 4.4 in AmyLQS45 or increased to 9.0 in AmyLQS46, when compared with the calculated pI of AmyL which was 6.3. However, the calculated pI value is only an indication of the overall charge of a protein and may not necessarily correspond to the actual pI, which must be determined experimentally.

In practice the most extreme forms of chimeric α-amylases, AmyLQS45 and AmyLQS46, would have been time consuming to construct using recombinant DNA techniques because of the need to replace numerous small fragments of DNA and the absence of suitable restriction sites. It was therefore decided to construct the intermediate chimeric α-amylases, AmyLQS55 and AmyLQS50, because of their ease of construction and relatively marked changes in calculated pI (4.9 and 8.1, respectively).

The chimeric α-amylases were designed to preserve the three dimensional structure and therefore, the catalytic activities of the enzymes. However, the effects of individual amino acid changes on the activities of the chimeric enzymes were not predictable from the protein sequence data particularly in the absence of the crystal structure of AmyL at the start of the project. Table 5.2 indicates the numbers of charged amino acids in wild type AmyL and in the chimeric α-amylases AmyLQS50 and AmyLQS55. The calculated pI values of the proteins are also shown.

Table 5.2. The total number of charged amino acids and the calculated pI values of the mature forms of AmyL, AmyLQS50 and AmyLQS55

	AmyL	AmyLQS50	AmyLQS55
Positively charged amino acids	73	82	59
Negatively charged amino acids	62	60	68
Calculated pI value	6.3	8.1	4.9

Positively charged amino acids - lysine, arginine, histidine

Negatively charged amino acids - aspartic acid, glutamic acid

AmyLQS50 was constructed by replacing various regions of AmyL with corresponding regions of AmyQ and AmyS. This increased the number of positively charged amino acids from 73 to 82, and decreased the number negatively charged amino acids from 62 to 60. AmyLQS55 was constructed in an analogous manner, by reducing the number of positively charged amino acids from 73 to 59 and increasing the number of negatively charged amino acids from 62 to 68. Therefore, work was initiated to construct the genes encoding chimeras AmyLQS50 and AmyLQS55. The gene encoding the low pI protein, AmyLQS55, was constructed by C L Jensen and the discussion here will concentrate on the construction of the gene encoding the high pI α-amylase AmyLQS50, and intermediates with an increased positive charge. A sequence alignment of AmyL and AmyLQS50.5 is shown in Fig. 5.2.

The mature AmyS protein is considerably larger than AmyL and AmyQ (Table 5.1 and Fig. 5.1). The extra length of AmyS is due to the presence of a long carboxy-terminal "tail" which does not appear to be important for the catalytic activity of this enzyme (Gray *et al.*, 1986). Since a large number of the amino acid residues making up the tail are positively charged, this region was used to replace corresponding amino acids from AmyL during the design of AmyLQS50 (Fig. 5.2). Incorporation of the AmyS tail into chimeric α-amylases allowed a considerable increase in calculated pI.

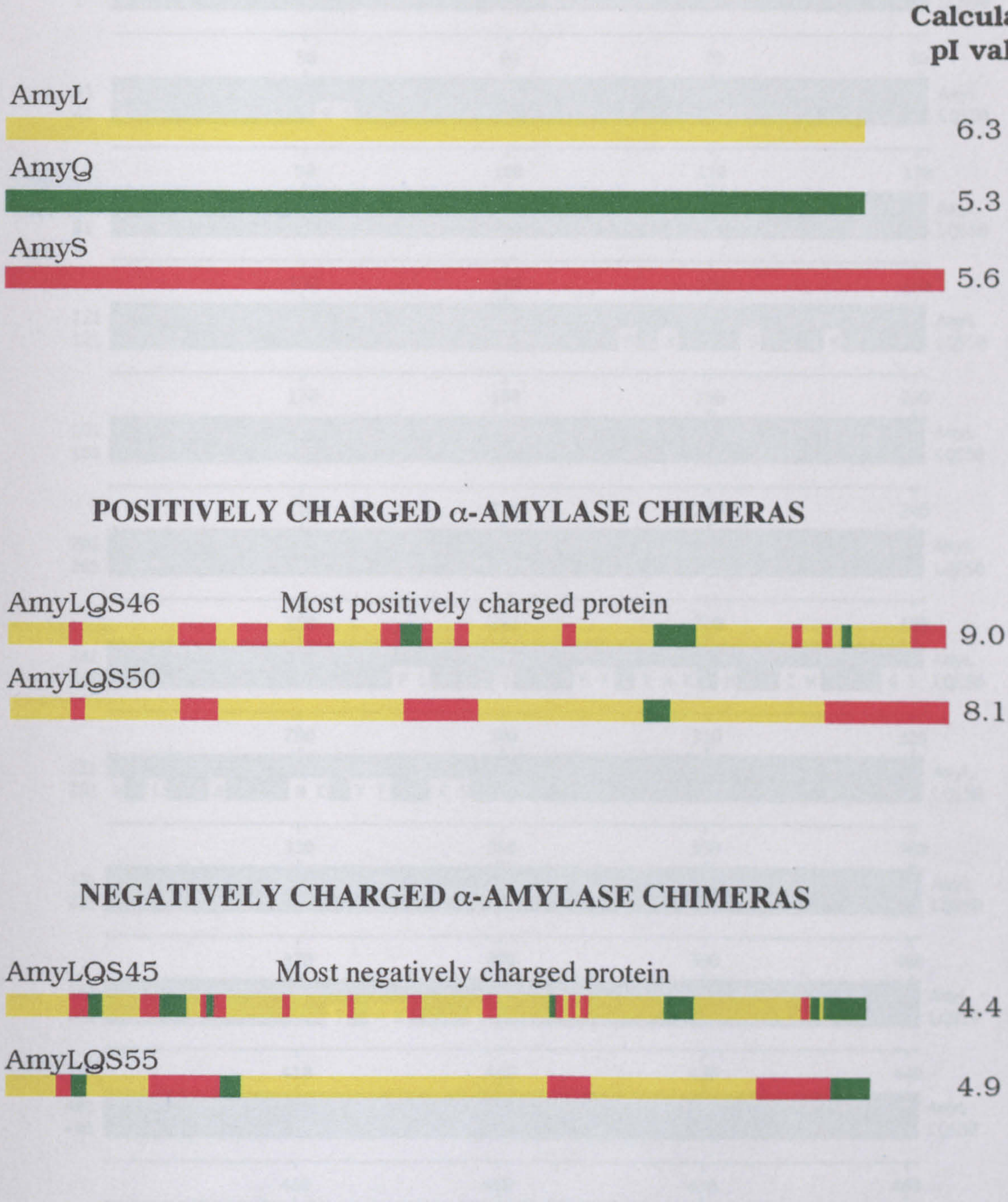


Fig. 5.1. Schematic representation and calculated pI of the mature regions of the three related *Bacillus* α-amylases together with *in silico* generated chimeric α-amylases.

Fig. 5.2. Sequence alignment of AmyL and AmyLQS50.
Residues identical to AmyL are shaded in black.

5.3 Construction of the genes encoding chimeric α-amylases

All of the chimeric α-amylase genes were to be expressed in *B. subtilis* using the xylose-inducible system (2.19) and exported using the AmyL signal peptide. The restriction maps of *amyL* and *amyLQS50* are shown in Fig. 5.3 and the sequences of the genes can be found in Appendix 1. The signal peptidase cleavage site of AmyL lies downstream of the *Pst*I site and the transcription terminator is centered 235 bp upstream of the *Hind*III site in an untranslated region.

In order to use the signal sequence and transcriptional terminator of the *amyL* gene, the chimeric α-amylase genes were constructed as in-frame *Pst*I to *Hind*III DNA fragments encoding the mature regions only. *amyLQS50* was constructed from five "blocks" of DNA in which five regions of *amyL* were replaced by corresponding regions of either *amyQ* or *amyS* (Table 5.3). The presence of DNA encoding the AmyS "tail" increased the size of the *amyLQS50* *Pst*I to *Hind*III fragment to 1.820 kbp, as compared with the corresponding fragment from *amyL* (1.723 kbp)

Table 5.3. The nucleotide and corresponding amino acid replacements during the construction of *amyLQS50*

Wild type gene/protein		Chimeric gene/protein			
amyL	AmyL	*amyQ*	AmyQ	*amyS*	AmyS
Nucleotides	Amino acids	Nucleotides	Amino acids	Nucleotides	Amino acids
232 - 255	49 - 56	-	-	250 - 273	50 - 57
517 - 558	144 - 157	-	-	535 - 576	145 - 158
841 - 981	252 - 298	-	-	865 - 1005	255 - 301
1198 - 1227	371 - 380	1204 - 1233	372 - 381	-	-
1399 - 1536	438 - 483	-	-	1414 - 1647	438 - 515

The distances (bp) were calculated in relation to the first nucleotide of the start codon. The transcription terminators of the *amyL*, *amyQ* and *amyS* genes are located at positions 1537 - 1539, 1543 - 1545 and 1648 - 1650, respectively. The amino acid positions were calculated from the start of the mature proteins and not the translated sequences.

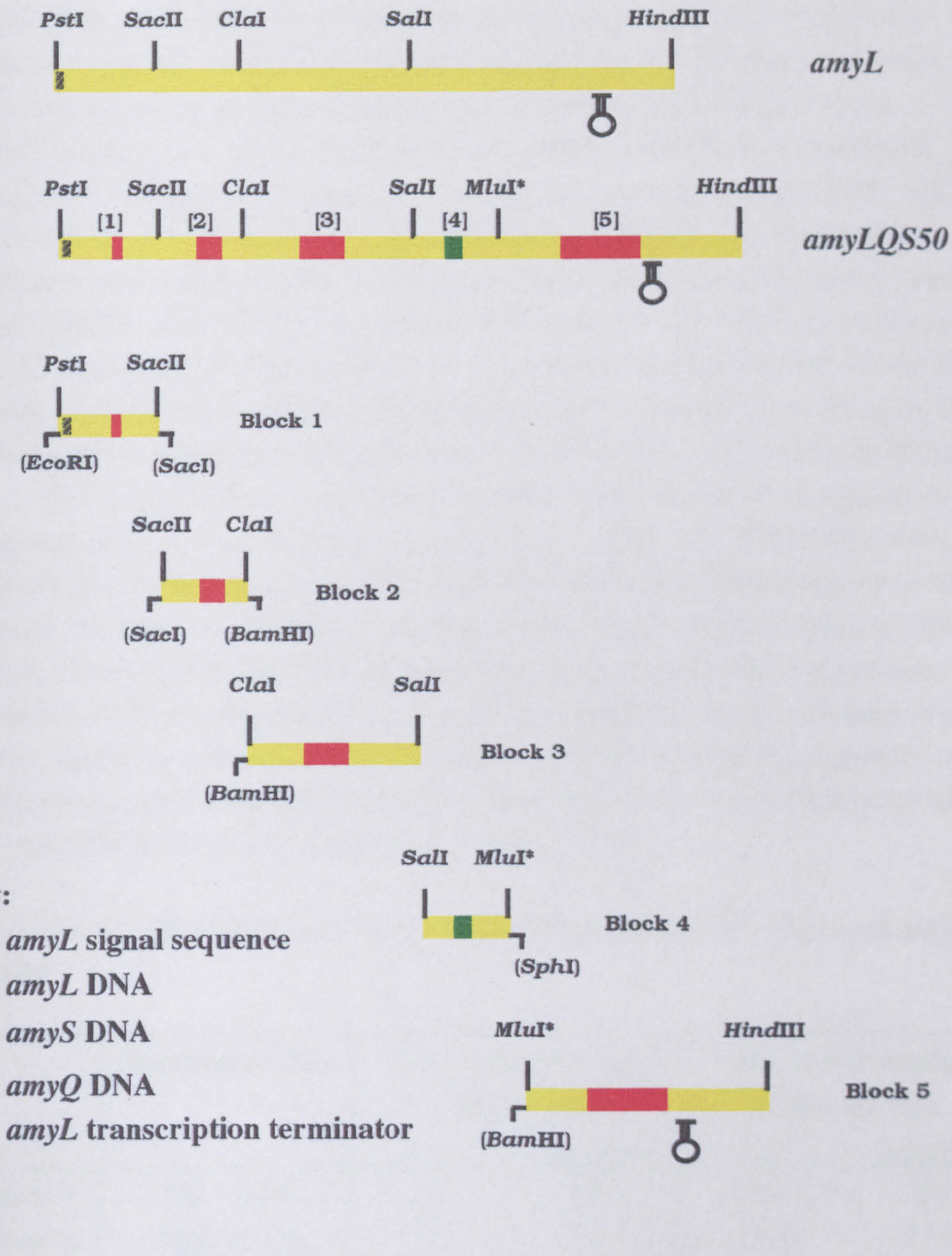


Fig. 5.3. Schematic representation of the restriction maps of *amyL* and *amyLQS50* and also the component blocks making up the *amyLQS0* gene.

The restriction sites shown above the gene/blocks are those within *amyL*,except in the case of *Mlu*I* which was created by the introduction of a single silent mutation. Restriction sites in brackets were those created during PCR to facilitate the cloning of blocks. Block numbers are indicated in square brackets.

5.3.1 Strategy for the construction of the *amyLQS50* gene

The construction of the gene encoding the chimeric α-amylase AmyLQS50 was carried out in several stages and is discussed in detail in the following sections. To facilitate the exchange of five specific regions of *amyL* with corresponding regions of *amyQ* or *amyS* (Table 5.3), *amyLQS50* was assembled from five separate blocks of DNA, each of which was constructed using the PCR-based technique, gene splicing by overlap extension (SOE, 5.3.2). The location of each block was determined by the presence of several unique restriction sites within the *amyL* gene and, in all cases, the total block size was larger than the relative restriction fragment (Table 5.4). The *Mlu*I site shown in Fig. 5.3 was not present in *amyL* but was engineered into the relevant blocks by creating a single silent mutation. Individual blocks were constructed using SOE, cloned into the multiple cloning site (MCS) of pUC19 in *E. coli* and sequenced.

To facilitate cloning and assembly, some of the blocks were engineered with additional restriction sites at one or both ends (Fig. 5.3). These sites were built into the blocks as required during the SOE PCR using flanking primers with relevant 5' extensions. Following sequence verification, the blocks were used to assemble *amyLQS50* in pUC19. As a control the *amyL* gene was reconstructed using five PCR-generated wild type blocks. The amplification of wild type *amyL* blocks equivalent to the chimeric blocks of *amyLQS50* allowed the assembly of intermediate genes with different combinations of chimeric and wild type blocks and encoding proteins with different pI values (5.3.7b).

Table 5.4. Constructed blocks used for the assembly of *amyLQS50* and *amyL* control

	Restriction sites	Size of constructed block (bp)		Size of restriction fragment (bp)	
		amyL	*amyLQS50*	*amyL*	*amyLQS50*
Block 1	*Pst*I - *Sac*II	405	405	362	362
Block 2	*Sac*II - *Cla*I	292	292	252	252
Block 3	*Cla*I - *Sal*I	443	443	407	407
Block 4	*Sal*I - *Mlu*I*	215	215	172	172
Block 5	*Mlu*I* - *Hind*III	563	659	530	627
Total	*Pst*I - *Hind*III	-	-	1723	1820

An asterix indicates a restriction site created by silent mutation

5.3.2 Gene splicing by overlap extension

The method chosen for the construction of *amyLQS50* was SOE, a PCR-based system (Horton *et al.*, 1989). This technique allows regions of DNA to be spliced at specific locations without the use of restriction enzymes and DNA ligase, thus facilitating the construction of defined gene fusions in the absence of suitable restriction sites. A schematic representation of SOE is shown in Fig. 5.4 and the basic principles are discussed below.

In order to splice two regions (I and II) of DNA together four oligonucleotide primers (-a to -d) are required. The two primers at the fusion junction, primer-b and primer-c, carry extensions at their 5' ends which are complementary to the 3' portion of the other primer, primer-c or primer-b, respectively. The flanking primers (primer-a and primer-d) are complementary to their respective template DNA only. In the initial rounds of PCR, primer-a and primer-b are used to amplify region I (product-ab) in PCR I and simultaneously, primer-c and primer-d are used to amplify region II (product-cd) in PCR II. As a consequence of the complementary sequences in the ends of primer-b and primer-c, product-ab and product-cd also have complementary sequences at one end of each product corresponding to the fusion junction.

The splicing of these two products occurs in the final reaction, PCR III, which proceeds by a two-step mechanism. In this reaction, product-ab, product-cd, primer-a and primer-d are mixed together and subjected to repeated rounds of denaturation, annealing and extension by *Taq*. Since one strand of each PCR product has complementary sequences in their 3' ends, they are able to serve as primers for each other in the first step of PCR III to yield product-ad, the complete region I/II spliced product. The second step of PCR III involves the amplification of the newly synthesized product-ad using primer-a and primer-d. The strands of product-ab and product-cd with the complementary sequences in their 5' ends are not capable of the self-priming step and therefore, do not contribute to the synthesis of the recombinant product. If required, the flanking primers can be designed with 5' extensions incorporating suitable restriction sites to facilitate subsequent cloning of the spliced product. SOE was used to construct the chimeric α-amylase gene *amyLQS50* (and *amy-LQS55*) by splicing specific regions of the genes encoding the three related *Bacillus* α-amylases.

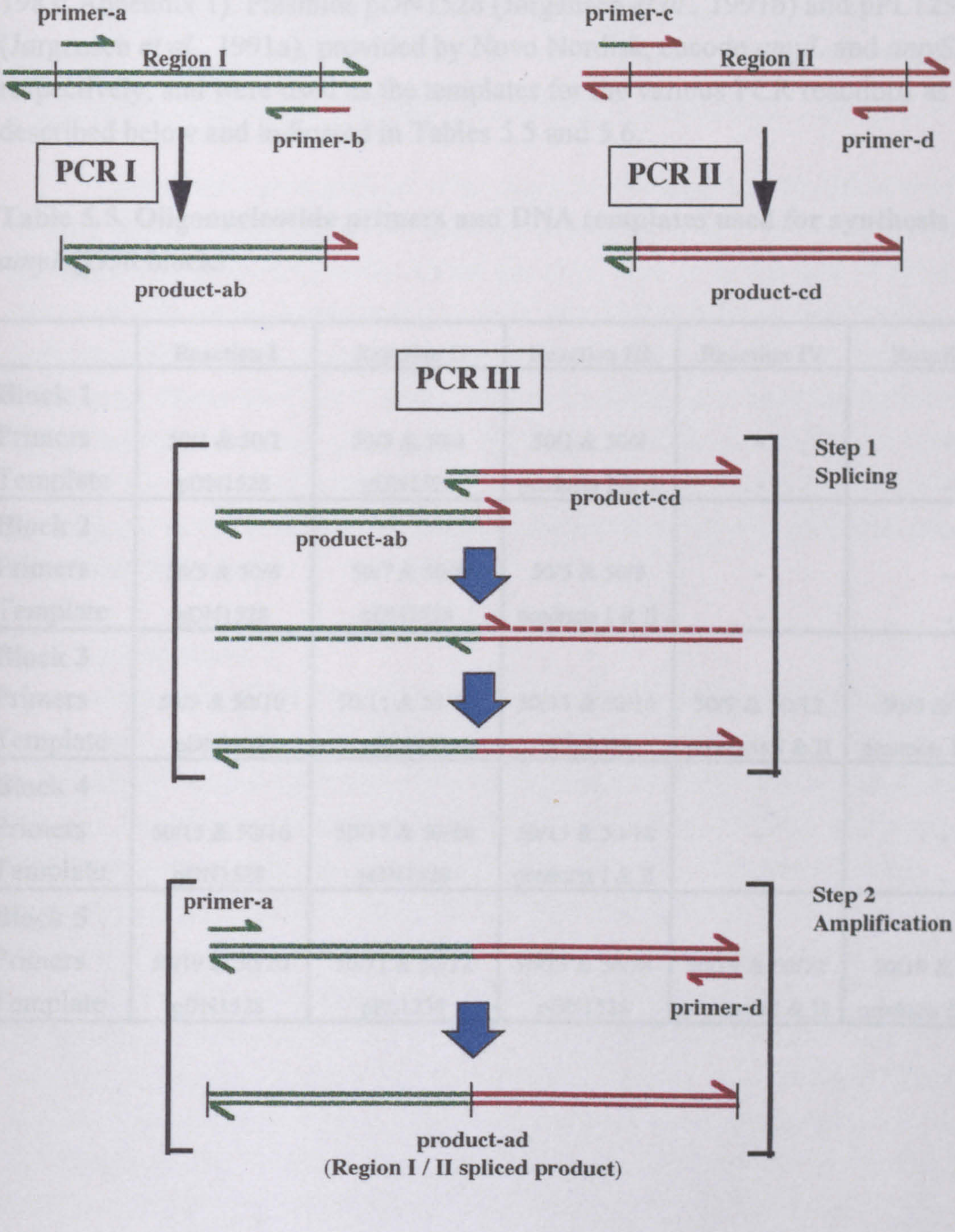


Fig. 5.4. Schematic representation of gene splicing by overlap extension.

The 3' ends of the DNA strands are indicated by arrows.

5.3.3 Synthesis of the component blocks of *amyLQS50* and *amyL*

The genes encoding the three related *Bacillus* α-amylases, AmyL, AmyS and AmyQ, have been cloned and sequenced (Gray *et al.*, 1986; Takkinen *et al.*, 1983; Appendix 1). Plasmids pDN1528 (Jørgensen *et al.*, 1991b) and pPL1235 (Jørgensen *et al.*, 1991a), provided by Novo Nordisk, encode *amyL* and *amyS*, respectively, and were used as the templates for the various PCR reactions as described below and indicated in Tables 5.5 and 5.6.

Table 5.5. Oligonucleotide primers and DNA templates used for synthesis of *amyLQS50* blocks

	Reaction I	Reaction II	Reaction III	Reaction IV	Reaction V
Block 1					
Primers	50/1 & 50/2	50/3 & 50/4	50/1 & 50/4	-	-
Template	pDN1528	pDN1528	products I & II	-	-
Block 2					
Primers	50/5 & 50/6	50/7 & 50/8	50/5 & 50/8	-	-
Template	pDN1528	pDN1528	products I & II	-	-
Block 3					
Primers	50/9 & 50/10	50/11 & 50/12	50/13 & 50/14	50/9 & 50/12	50/9 & 50/14
Template	pDN1528	pPL1235	pDN1528	products I & II	products III & IV
Block 4					
Primers	50/15 & 50/16	50/17 & 50/18	50/15 & 50/18	-	-
Template	pDN1528	pDN1528	products I & II	-	-
Block 5					
Primers	50/19 & 50/20	50/21 & 50/22	50/23 & 50/24	50/19 & 50/22	50/19 & 50/24
Template	pDN1528	pPL1235	pDN1528	products I & II	products III & IV

Continued

Table 5.6. Oligonucleotide primers and DNA templates used for synthesis of *amyL* blocks

	Block 1	Block 2	Block 3	Block 4	Block 5
Primers	50/1 - 50/4	50/5 - 50/8	50/9 - 50/14	50/15 - 50/18	50/19 - 50/24
Template	pDN1528	pDN1528	pDN1528	pDN1528	pDN1528

Schematic representations of the steps involved in the construction of the various blocks are shown below. In all cases restriction sites above the DNA represent sites present naturally within the *amyL* gene or in the case of *Mlu*I, a site created by silent mutation during PCR. Restriction sites in parentheses indicate artificial sites created during PCR by primers with the relevant sites as 5' extensions. These sites were incorporated to facilitate cloning of the blocks and the assembly of the *amyLQS50* gene. The locations of all nucleotides or amino acid residues described below were calculated with reference to the start codons of the genes or the first residue of the mature protein, respectively (consistent with Table 5.3).

5.3.3a Block 1

amyLQS50 block 1 (LQS1) was designed to facilitate the replacement of *amyL* nucleotides 232 to 255 (amino acids 49 to 56) with *amyS* nucleotides 250 to 273 (AmyS amino acids 50 to 57). Its construction is shown diagramatically in Fig. 5.5. As indicated in Table 5.5, four primers (primer-LQS50/1 to primer-LQS50/4) were required for the amplification reactions. The two flanking primers, primer-LQS50/1 (34-mer) and primer-LQS50/4 (34-mer), were designed to anneal to *amyL* over the *Pst*I and *Sac*II sites, respectively. Primer-LQS50/1 was designed with a 5' *Eco*RI site and primer-LQS50/4 with a 5' *Sac*I site, to facilitate the cloning of the PCR product. Primer-LQS50/2 was a 46-mer of which 22 nucleotides at the 3' end corresponded to *amyL* nucleotides 210 to 231 and the 24 nucleotides at the 5' end corresponded to *amyS* nucleotides 250 to 273. Similarly, the 3' end of primer-LQS50/3 annealed to *amyL* nucleotides 256 to 281 while the 5' nucleotides annealed to nucleotides 250 to 273 of *amyS*.

The PCR was carried out (3.12) using pDN1528 as the template. Reaction-I was carried out using primer-LQS50/1 and primer-LQS50/2 to generate product I (206 bp). Product II (223 bp) was amplified in Reaction-II using primer-LQS50/3 and primer-LQS50/4. The PCR reactions were analysed by agarose electrophoresis (3.6) and products I and II were of the expected size

(Fig. 5.6). The PCR products were then purified by extraction from agarose gels (3.7).

Purified PCR products I and II were used as template and primers for Reaction-III, which results in the splicing of the two products. This was facilitated by the complementary sequences (*amyS* nucleotides 250 to 273) in one end of product I and product II. The initial splicing reaction was followed immediately by an amplification reaction using primer-LQS50/1 and primer-LQS50/4. The final product, of 405 bp in length, was designated LQS1 and corresponded to the replacement of *amyL* nucleotides 232 to 255 with nucleotides 250 to 273 from *amyS*.

Plasmid pDN1528 DNA was subjected to PCR with primer-LQS50/1 and primer-LQS50/4 to provide a wild type *amyL* block 1 (WT1). The spliced and wild type products were analyzed by agarose electrophoresis and both fragments were of the expected size (Fig. 5.6). However, in addition to the block 1 products, there were a few minor products of smaller size present in the *amyLQS50* SOE reaction. The nature of these products was not investigated further and the WT1 and LQS1 fragments were gel purified.

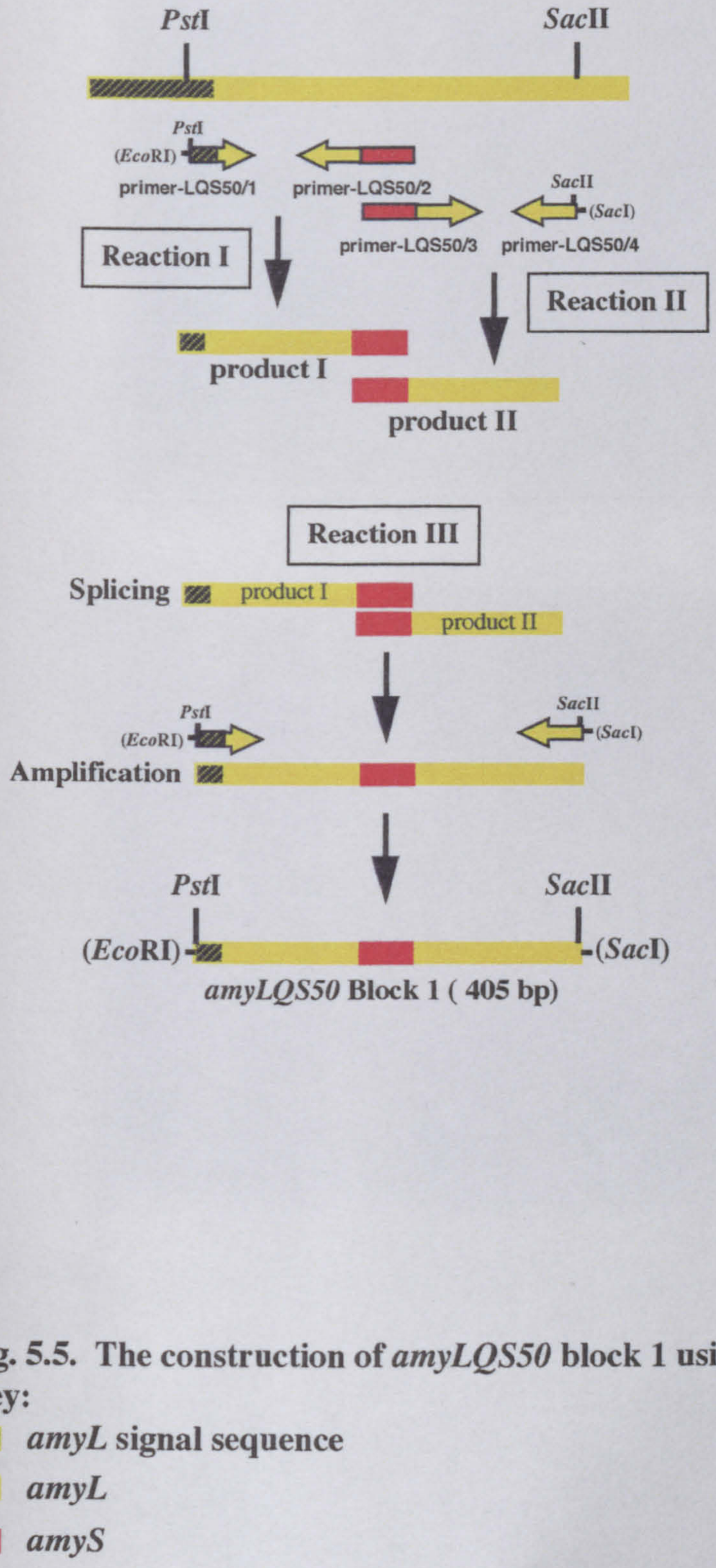


Fig. 5.5. The construction of *amyLQS50* block 1 using SOE.
Key:

- ▨ ***amyL* signal sequence**
- ▭ ***amyL***
- ▭ ***amyS***

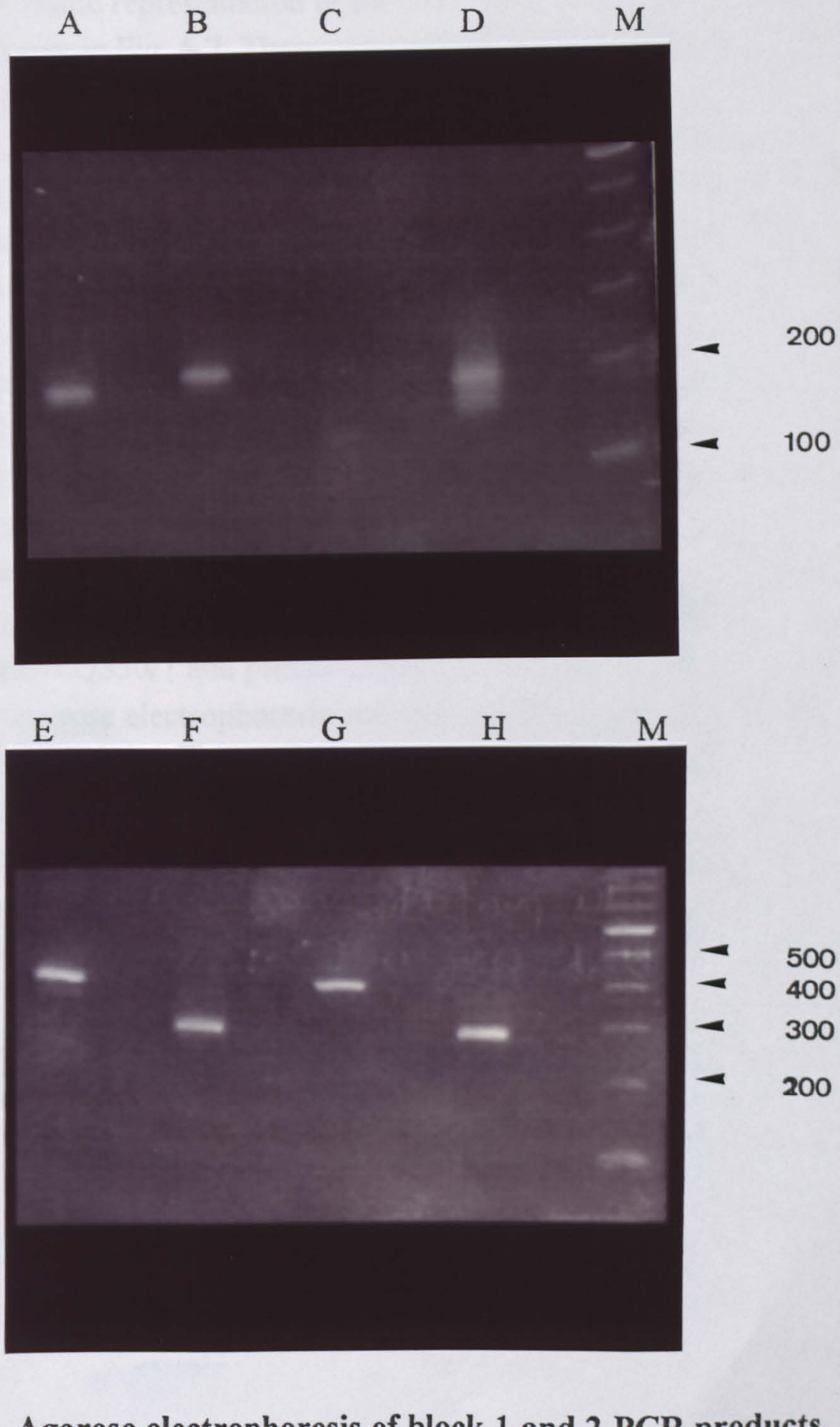


Fig. 5.6. Agarose electrophoresis of block 1 and 2 PCR products.

Key:

M - 100 bp ladder
A - Block 1 product-I
B - Block 1 product-II
C - Block 2 product-I
D - Block 2 product-II
E - LQS1
F - LQS2
G - WT1
H - WT2

5.3.3b Block 2

A schematic representation of the construction of *amyLQS50* block 2 (LQS2) is shown in Fig. 5.7. Three separate reactions and 4 primers were required to construct LQS2 which resulted in the replacement of *amyL* nucleotides 517 to 558 (amino acid residues 144 to 157) with *amyS* nucleotides 535 to 576 (amino acids 145 to 158). The sequences of the block 2 flanking primers, primer-LQS50/5 (34-mer) and primer-LQS50/8 (35-mer) were such that they annealed over the *Sac*II and *Cla*I restriction sites of *amyL*, respectively. In addition, primer-LQS50/5 and primer-LQS50/8 carried 5' extensions to introduce *Sac*I and *Bam*HI restriction sites into the ends of their respective PCR products. The internal primers, primer-LQS50/6 (62-mer) and primer-LQS50/7 (65-mer), were designed with 42 nucleotide 5' extensions which were complementary to nucleotides 535 to 576 of *amyS* (amino acids 145 to 158).

Plasmid pDN1528 was used as template for Reactions-I and -II. PCR product I was amplified using primer-LQS50/5 and primer-LQS50/6 and product II using primer-LQS50/7 and primer-LQS50/8. The generated PCR products were analyzed by agarose electrophoresis and shown to be of the expected sizes; 141 bp for product I and 196 bp for product II (Fig. 5.6). The PCR products were then purified from agarose slices following electrophoresis.

Purified block 2 products I and II were used as templates for the splicing reaction, Reaction-III, using the complementary sequences at the ends of the products. The spliced product was amplified using primer-LQS50/5 and primer-LQS50/8 to form the complete LQS2 fragment, in which nucleotides 517 to 558 of *amyL* were replaced by *amyS* nucleotides 535 to 576. The wild type block 2 fragment (WT2) was amplified from pDN1528 using primer-LQS50/5 and primer-LQS50/8. The WT2 and LQS2 fragments were run on an agarose gel (Fig. 5.6) to confirm that both were of the size estimated from the DNA sequence (294 bp). The products were then agarose gel purified.

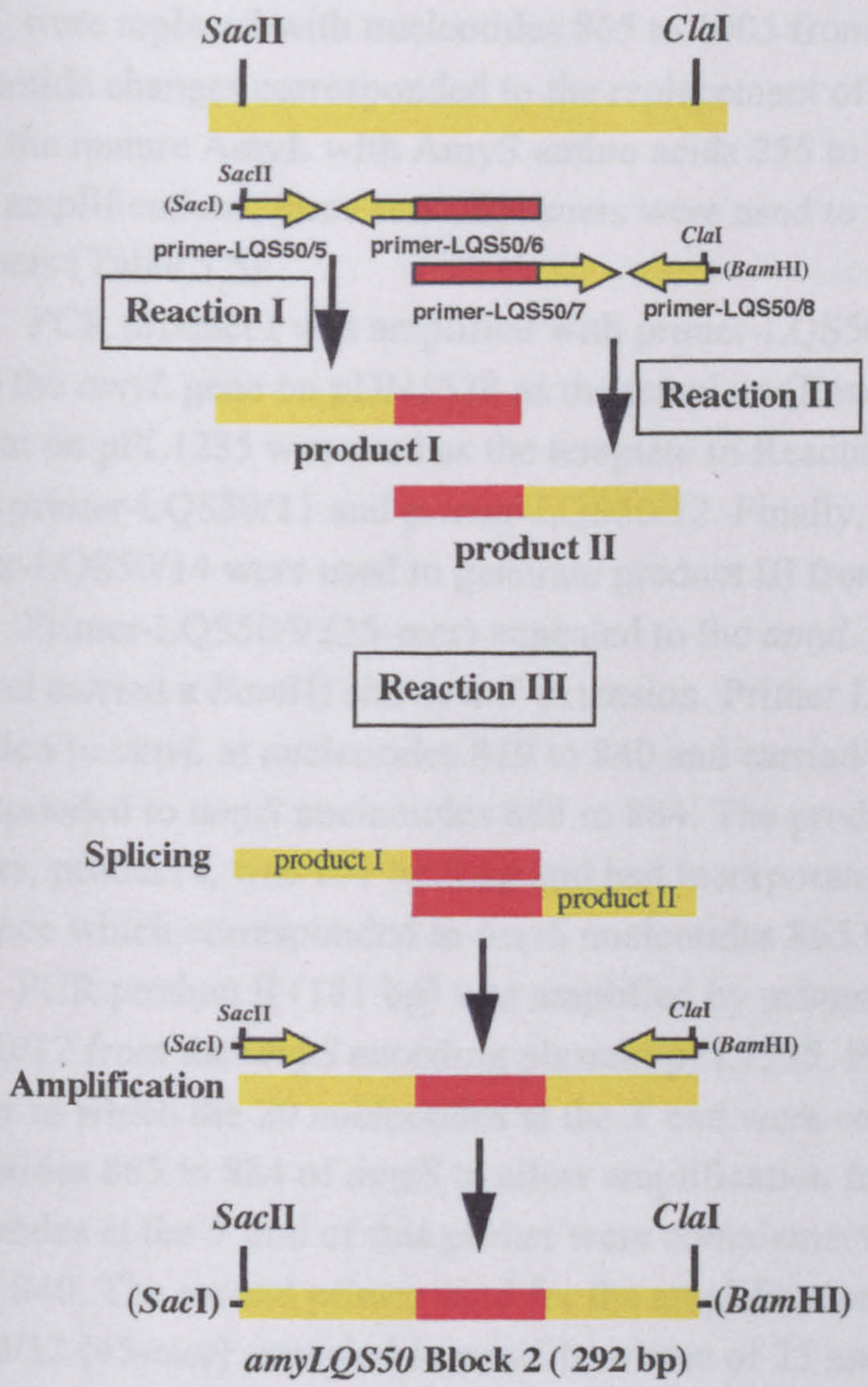


Fig. 5.7. The construction of *amyLQS50* block 2 using SOE.
Key:

amyL

amyS

5.3.3c Block 3

The construction of *amyLQS50* block 3 (LQS3) was more complicated than for previous blocks (Fig. 5.8) and required the use of 6 primers and 5 separate amplification reactions. To construct LQS3, nucleotides 841 to 981 from *amyL* were replaced with nucleotides 865 to 1005 from *amyS* using SOE. These nucleotide changes corresponded to the replacement of amino acids 252 to 298 from the mature AmyL with AmyS amino acids 255 to 301. In the initial round of PCR amplifications, three sets of primers were used to amplify three separate products (Table 5.5).

PCR product I was amplified with primer-LQS50/9 and primer-LQS50/10 using the *amyL* gene on pDN1528 as the template (Reaction-I). The *amyS* gene present on pPL1235 was used as the template in Reaction-II to amplify product II using primer-LQS50/11 and primer-LQS50/12. Finally, primer-LQS50/13 and primer-LQS50/14 were used to generate product III from the *amyL* template.

Primer-LQS50/9 (35-mer) annealed to the *amyL* template over the *Cla*I site and carried a *Bam*HI site as a 5' extension. Primer LQS50/10 (41-mer) annealed to *amyL* at nucleotides 819 to 840 and carried a 5' extension which corresponded to *amyS* nucleotides 865 to 884. The product generated by these primers, product I, was 191 bp long and had incorporated into one end a DNA sequence which corresponded to *amyS* nucleotides 865 to 884.

PCR product II (181 bp) was amplified by primer-LQS50/11 and primer-LQS50/12 from the *amyS* encoding plasmid pPL1235. Primer-LQS50/11 was a 40-mer in which the 20 nucleotides at the 3' end were complementary to nucleotides 865 to 884 of *amyS* to allow amplification from this template. The nucleotides at the 5' end of this primer were complementary to *amyL* nucleotides 821 to 840. The second primer used for the amplification of product II, primer-LQS50/12 (45-mer) annealed to *amyS* by virtue of 25 nucleotides in its 3' end (*amyS* nucleotides 881 to 1005). The nucleotides at its 5' end were complementary to *amyL* nucleotides 982 to 1001. Product II therefore consisted of nucleotides 865 to 1005 from *amyS* with extensions at both ends which were complementary to specific regions of *amyL*.

Product III (152 bp) was amplified from pDN1528 using the primer pair primer-LQS50/13 and primer-LQS50/14. Twenty nucleotides at the 3' end of primer-LQS50/13 (40-mer) annealed to *amyL* over nucleotides 982 to 1001. Primer-LQS50/14 (25-mer) annealed to the *amyL* template over the *Sal*I restriction site. Product III therefore consisted primarily of *amyL* sequences but with *amyS* nucleotides 985 to 1005 at one end. Following PCR the amplification products of Reactions-I, -II and -III were analysed by agarose electrophoresis and

shown to be of the expected size (Fig. 5.9). These PCR products were then gel purified.

PCR products I, II and III were designed with complementary sequences at one or both ends. Block 3 was assembled from these products by two additional SOE PCR stages. In the first stage the overlapping PCR products I and II, which have complementary ends, were used as template and primers for Reaction-IV. This reaction occurred by the two step-mechanism (5.3.2) with subsequent amplification of product IV (323 bp) with primer-LQS50/9 and primer-LQS50/12. The size of product IV was verified by agarose electrophoresis (Fig. 5.9) and the fragment was gel purified.

The final SOE stage in the construction of block 3 involved the splicing of overlapping products IV and III (Reaction-V) with amplification of the spliced product by primer-LQS50/9 and primer-LQS50/14. The size of the product generated, LQS3, was confirmed by agarose electrophoresis at 443 bp (Fig. 5.9). In addition to the *Cla*I and *Sal*I restriction sites from the *amyL* gene, LQS3 has a flanking *Bam*HI site at the *Cla*I end which was derived from primer-LQS50/9.

Plasmid pDN1528 was used as template to amplify the *amyL* wild type block 3 (WT3), using primer pair primer-LQS50/9 and primer-LQS50/14. The size of this product was also verified by electrophoresis (Fig. 5.9) and both WT3 and LQS3 fragments were gel purified.

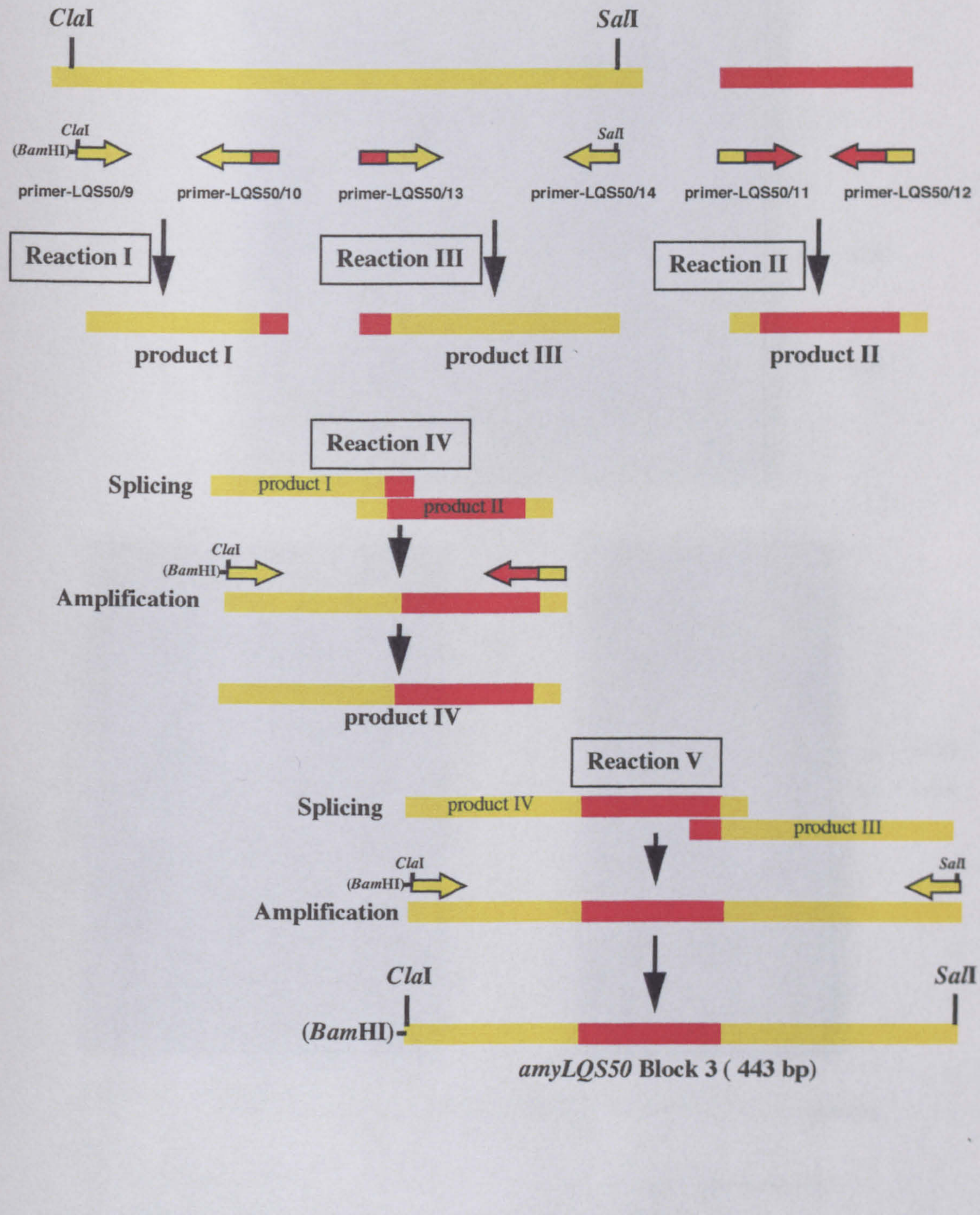


Fig. 5.8. The construction of *amyLQS50* block 3 using SOE.
Key:

amyL

amyS

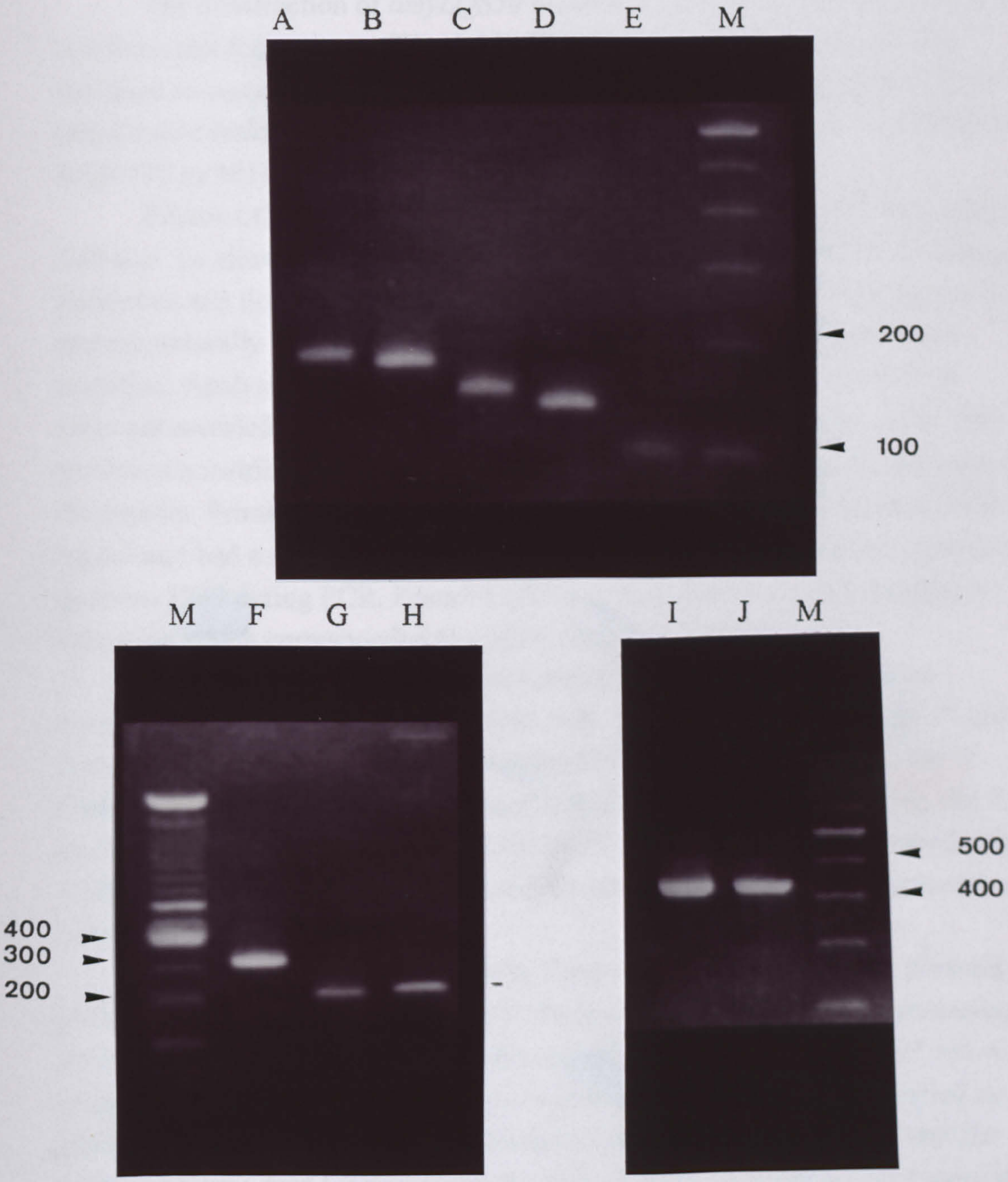


Fig. 5.9. Agarose electrophoresis of block 3 and 4 PCR products.

Key:

M - 100 bp ladder
A - Block 3 product-I
B - Block 3 product-II
C - Block 3 product-III
D - Block 4 product-I
E - Block 4 product-II
F - Block 3 product-IV
G - LQS4
H - WT4
I - LQS3
J - WT3

5.3.3d Block 4

The construction of *amyLQS50* block 4 (LQS4) used three SOE PCR reactions and four primers (Fig. 5.10). The DNA sequence of block 4 was designed to correspond to the replacement of *amyL* nucleotides 1198 to 1227 (amino acid residues 371 to 380) with *amyQ* nucleotides 1204 to 1233 (amino acids 372 to 381).

Primer-LQS50/15 (30-mer) annealed to the *amyL* gene over the unique *Sal*I site. To allow the replacement of *amyL* nucleotides 1198 to 1227, a unique restriction site downstream of this region was required. Since no suitable site was present naturally in the *amyL* gene it was decided to create a site by silent mutation. Analysis of the sequence of the *amyL* gene using the PC/GENE software revealed that a T to C nucleotide change at nucleotide position 1269 produced a restriction site for *Mlu*I but did not alter the amino acid sequence of the protein. Primer-LQS50/18 (39-mer) was designed to anneal to *amyL* over this region and had a single nucleotide mismatch to introduce the desired mutation at position 1269 during PCR. Primer-LQS50/18 also carried a non-annealing 5' extension which corresponded to a *Sph*I site.

Primer-LQS50/16 (38-mer) and primer-LQS50/17 (43-mer) were complementary to *amyL* in their 3' ends only. The 20 nucleotides at the 3' end of primer-LQS50/16 annealed to nucleotides 1177 to 1197 of *amyL* and the 5' nucleotides corresponded to *amyQ* nucleotides 1204 to 1221. Similarly, the 13 nucleotides in the 5' end of primer-LQS50/17 annealed to *amyL* over nucleotides 1228 to 1240 and the twenty 5' nucleotides corresponded to *amyQ* nucleotides 1204 to 1233.

Two overlapping PCR products, I and II, were amplified from plasmid pDN1528 in the initial rounds of PCR. Product I (136 bp) was generated using primer pair primer-LQS50/15 and primer-LQS50/16 and product II (97 bp) with primer-LQS50/17 and primer-LQS50/18 and their respective sizes verified by agarose electrophoresis (Fig. 5.9). Products I and II were gel purified and the complementary *amyQ* sequences in the ends of these products allowed them to be spliced together in Reaction-III with subsequent amplification by the flanking primers, primer-LQS50/15 and primer-LQS50/18.

The wild type *amyL* block 4 (WT4) fragment was generated by PCR from the pDN1528 template with primer pair primer-LQS50/15 and primer-LQS50/18. The sizes of both WT4 (186 bp) and LQS4 (186 bp) fragments were analyzed by electrophoresis (Fig. 5.9). The WT4 fragment was of the size expected however, LQS4 appeared to be smaller than 186 bp. In spite of the apparent size difference of LQS4, the block 4 fragments were gel purified as before.

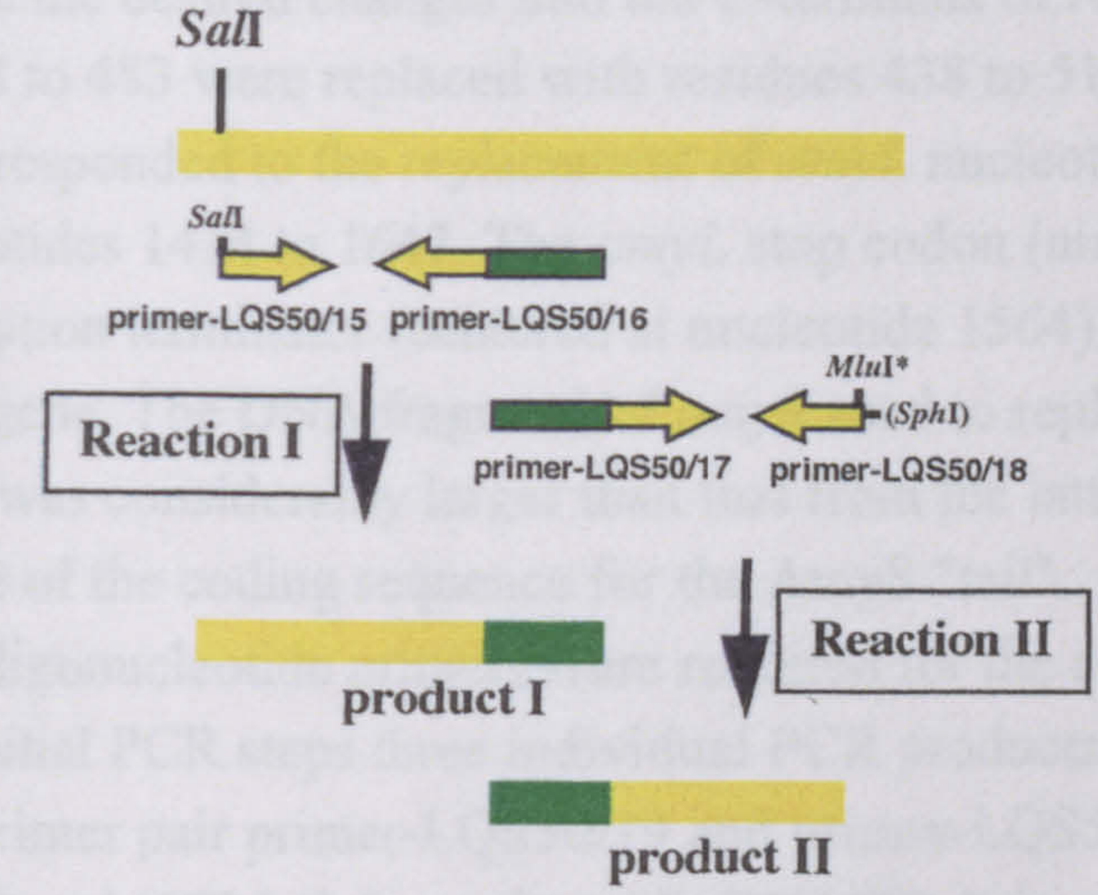


Fig. 5.10. The construction of *amyLQS50* block 4 using SOE.
Key:
amyL
amyS

5.3.3e Block 5

amyLQS50 block 5 (LQS5) was the largest of all the blocks and was constructed using five separate PCR steps in a similar way to LQS3 (Fig. 5.11). To introduce the desired changes into the C-terminus of AmyL, amino acid residues 438 to 483 were replaced with residues 438 to 515 from AmyS. These changes corresponded to the replacement of *amyL* nucleotides 1399 to 1536 with *amyS* nucleotides 1414 to 1647. The *amyL* stop codon (nucleotides 1537 to 1539) and transcription terminator (centered at nucleotide 1564) were used in the *amyLQS50* gene. The DNA fragment of *amyS* used to replace the corresponding *amyL* DNA was considerably larger than that from the latter gene as a consequence of the coding sequence for the AmyS "tail".

Six oligonucleotide primers were required for the construction of LQS5 and in the initial PCR steps three individual PCR products (I, II and III) were amplified. Primer pair primer-LQS50/19 and primer-LQS50/20 were used to amplify product I (173 bp) from plasmid pDN1528. Primer-LQS50/19 (38-mer) carried a T to C nucleotide mismatch at nucleotide 1269 to create the *Mlu*I site by silent mutation and a 5' *Bam*HI site to facilitate cloning. Primer-LQS50/20 (40-mer) was complementary to *amyL* in the twenty 3' nucleotides only, which corresponded to *amyL* nucleotides 1379 to 1398. The remaining nucleotides of this primer did not anneal to *amyL* but corresponded to *amyS* nucleotides 1414 to 1433.

Product II (233 bp) was amplified in Reaction-II using primer-LQS50/21 (41-mer) and primer-LQS50/22 (40-mer) from the *amyS* gene encoded by pPL1235. The twenty one nucleotides at the 3' end of primer-LQS50/21 were complementary to *amyS* nucleotides 1414 to 1434. The twenty nucleotides in the 5' end of the primer were complementary to nucleotides 1379 to 1398 from the *amyL* gene. Similarly, primer-LQS50/22 annealed to *amyS* by virtue of the twenty nucleotides in its 3' end which were complementary to *amyS* nucleotides 1628 to 1647. The twenty 5' nucleotides were complementary to amyL nucleotides 1537 to 1556.

Product III (284 bp) was amplified from plasmid pDN1528 using primer-LQS50/23 (40-mer) and primer-LQS50/24 (26-mer). The twenty nucleotides in the 3' end of primer-LQS50/23 were complementary to *amyL* nucleotides 1537 to 1556 and the twenty in the 5' end were complementary to *amyS* nucleotides 1628 to 1647. The primer at the end of *amyL* gene, primer-LQS50/24, was a 26-mer which was entirely complementary to this template and annealed over the flanking *Hind*III site.

Products I, II and III were amplified and the size of the fragments confirmed by agarose electrophoresis (Fig. 5.12). The fragments were purified

from agarose slices following electrophoresis and joined to form the block 5 fragment in two further SOE PCR stages. In the first stage, overlapping products I and II were spliced in Reaction IV to generate product IV (407 bp) with amplification by primer-LQS50/19 and primer-LQS50/22. The size of product IV was verified by electrophoresis (Fig. 5.12) and the fragment was gel purified.

In the last stage of the construction of LQS5, purified products IV and III were spliced and amplified by primer-LQS50/19 and primer-LQS50/24 to generate the complete block 5 fragment (659 bp). The *amyL* block 5 (WT5) fragment (563 bp) was generated by PCR from the plasmid pDN1528 template using the block 5 flanking primers, primer-LQS50/19 and primer-LQS50/24. Following the verification of fragment size by electrophoresis (Fig. 5.12) the WT5 and LQS5 fragments were gel purified.

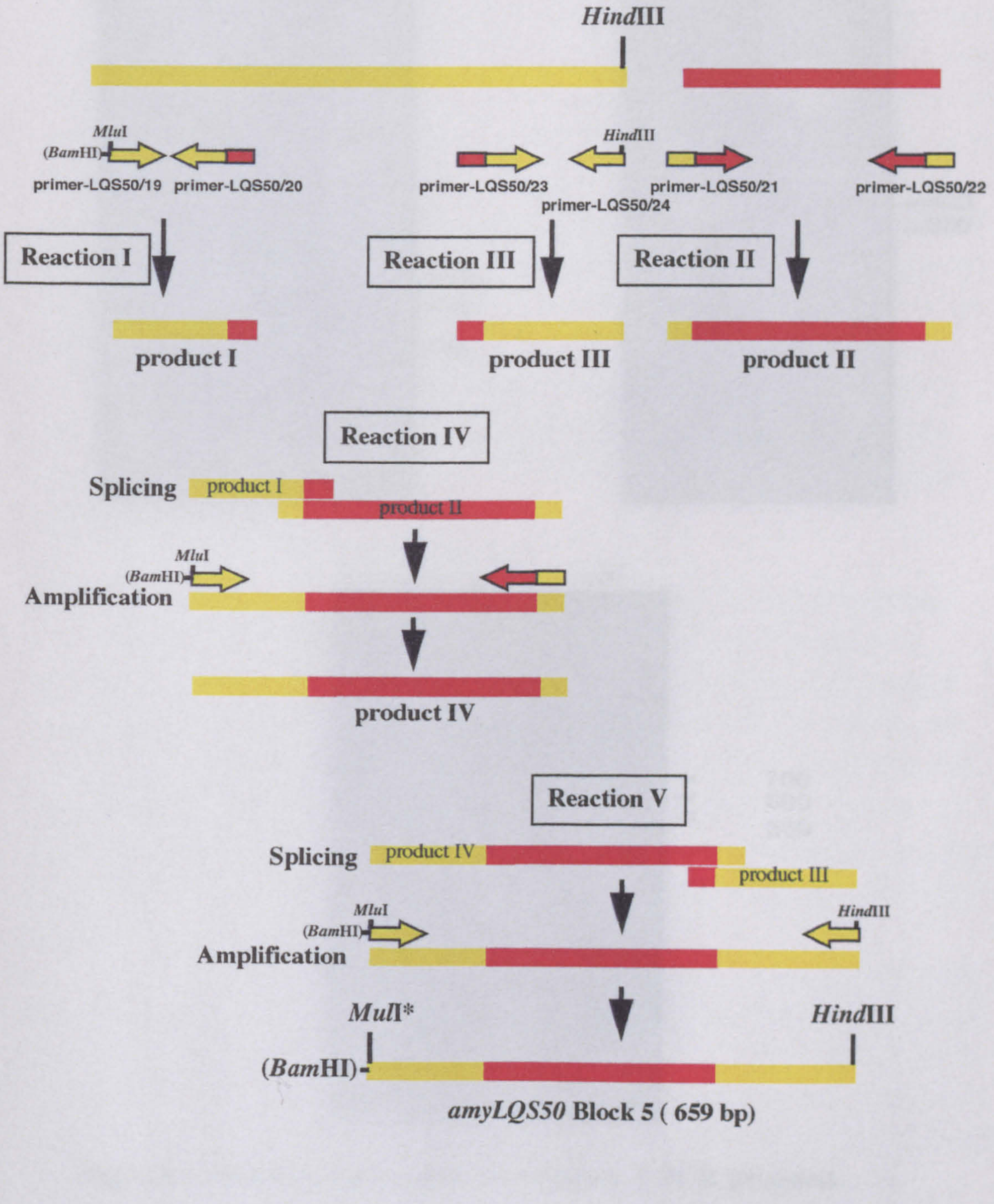


Fig. 5.11. The construction of *amyLQS50* block 5 using SOE.

amyL

amyS

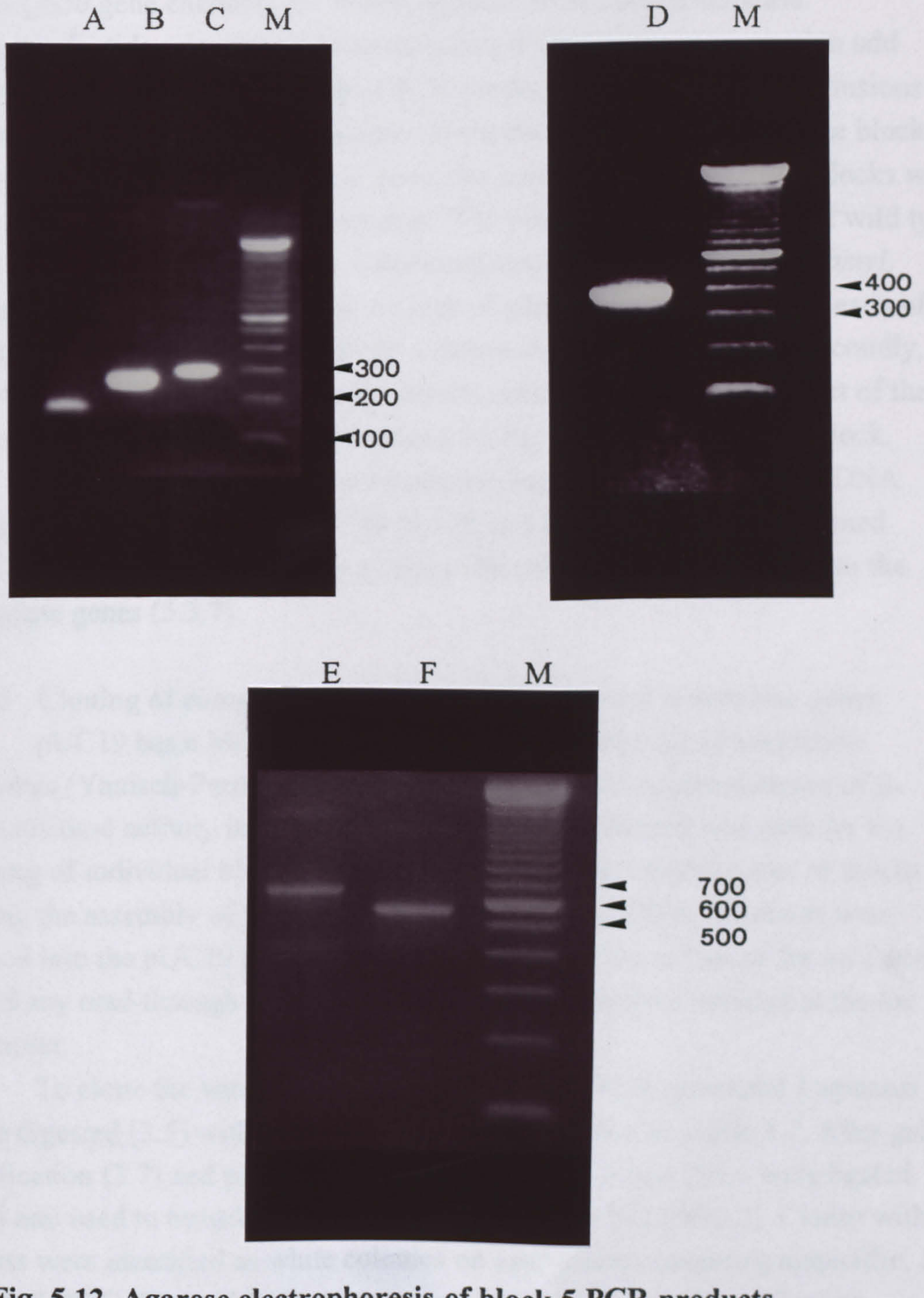


Fig. 5.12. Agarose electrophoresis of block 5 PCR products.

Key:

M - 100 bp ladder
A - Block 5 product-I
B - Block 5 product-II
C - Block 5 product-III
D - Block 5 product-IV
E - LQS5
F - WT5

5.3.4 Summary of SOE PCR

SOE PCR was utilized to construct the five component blocks for the *amyLQS50* gene encoding the mature form of a chimeric α-amylase. Oligonucleotide primers with non-annealing 5' extensions were used to add specific sequences into the ends of PCR products to allow exact gene fusions or permit the cloning of PCR fragments. Since the primers flanking all the blocks were complementary to the *amyL* gene, the corresponding wild type blocks were also amplified from the *amyL* template. The purpose of amplifying the wild type blocks was two fold. Firstly, by combining *amyLQS50* and wild type *amyL* blocks in different combinations, a range of genes could be produced, each of which would be expected to encode α-amylases with different pI's. Secondly, if one or more of the chimeric blocks proved unsuitable to be used as part of the *amyLQS50* gene, they could be replaced by the equivalent wild type block.

To ensure that the various blocks had been fused correctly, the DNA fragments were cloned into pUC19 (5.3.5) and their sequences determined (5.3.6). This was also necessary to allow the assembly of the blocks into the complete genes (5.3.7).

5.3.5 Cloning of component fragments of engineered α-amylase genes

pUC19 has a MCS with unique recognition sites for 13 restriction enzymes (Yanisch-Perron *et al.*, 1985) and allows α-complementation of β-galactosidase activity in suitable hosts (3.11). This plasmid was used for the cloning of individual blocks and also for constructing combinations of blocks during the assembly of various α-amylase genes. The DNA fragments were cloned into the pUC19 MCS in the opposite orientation to that of the *lacZ* gene to avoid any read-through of the insert DNA from transcripts initiated at the *lac* promoter.

To clone the various blocks, pUC19 and the PCR-generated fragments were digested (3.5) with the restriction enzymes shown in Table 5.7. After gel purification (3.7) and precipitation (3.3), vector and insert DNA were ligated (3.8) and used to transform electrocompetent *E. coli* SJ2 (3.10.2). Clones with inserts were identified as white colonies on agar plates containing ampicillin, X-Gal and IPTG. Plasmid DNA was extracted (3.4.2) from at least 12 white colonies arising from each transformation and the presence of inserts of the correct size verified by restriction digestion followed by electrophoresis.

Continued

Table 5.7. The restriction enzyme sites used to clone the various *amyLQS50* and *amyL* blocks into pUC19

	Restriction enzyme sites
Block 1	*Eco*RI § and *Sac*I §
Block 2	*Sac*I § and *Bam*HI §
Block 3	*Bam*HI § and *Sal*I
Block 4	*Sal*I and *Sph*I §
Block 5	*Bam*HI § and *Hind*III

§ Indicates a restriction site not present in *amyL* but created during PCR

Almost all of the white colonies tested contained inserts of the expected size, and on this basis several clones for each *amyLQS50* or wild type block were selected for DNA sequencing. To distinguish the various recombinant plasmids containing different blocks, a system of plasmid nomenclature was devised. After the lower case "p", the letters in upper case designate the block type, either LQS (for *amyLQS50* blocks) or WT (for wild type *amyL* blocks). The letters indicating block type are followed by a number (1-5) which relates to the block number. Finally, the individual clone numbers are indicated after a hyphen. For example, pLQS1-2 is a recombinant plasmid encoding *amyLQS50* block 1 which was isolated from clone number 2.

5.3.6 DNA sequencing of constructed blocks of *amyLQS50* and *amyL*

The DNA sequences of the various *amyLQS50* and wild type *amyL* blocks cloned into pUC19 were determined by sequencing across the MCS. Sequencing was carried out to confirm that the splicing and amplification reactions (5.3.3) had generated the expected products and to detect errors that may have resulted from the use of *Taq* polymerase, which does not exhibit proof-reading activity. Sequencing was carried out using the Sequenase kit (3.13) on plasmid DNA extracted using acid phenol (3.4.2).

At least two independent plasmid clones of each block were sequenced (Table 5.8) and the initial sequencing reactions were carried out using the universal M13 forward and reverse primers (Appendix 2). With the exception of block 5, the sizes of the blocks were less than 500 bp. Consequently, using short and long sequencing runs, it was possible to determine the sequence of the blocks with the universal primers. The LQS5 and WT5 fragments (659 and 563 bp,

respectively) were sequenced using the universal primers and additional internal oligonucleotides, primer-LQS50/25 and primer-LQS50/26 (Appendix 2).

Table 5.8. The plasmids sequenced during this study

	Plasmid	
	amyLQS50	*amyL*
Block 1	pLQS1-2	pWT1-11
	pLQS1-5	pWT1-12
	pLQS1-7	
Block 2	pLQS2-3	pWT2-8
	pLQS2-4	pWT2-9
Block 3	pLQS3-5	pWT3-15
	pLQS3-6	pWT3-16
Block 4	pLQS4-5	pWT4-18
	pLQS4-7	pWT4-20
	pLQS4-10	
Block 5	pLQS5-8	pWT5-11
	pLQS5-25	pWT5-17
	pLQS5-33	pWT5-18
	pLQS5-39	pWT5-19

The expected sequences of the various *amyLQS50* and wild type *amyL* blocks, as determined from the published sequences of *amyL*, *amyQ* and *amyS* (Gray *et al.*, 1986; Takkinen *et al.*, 1983), are shown in Figs. 5.13 to 5.17. As discussed in the following sections, a number of nucleotides were identified which did not correspond to the expected DNA sequence and these nucleotide changes were given specific designations (Table 5.9).

Continued

Table 5.9. Definition of the terms used to describe nucleotide changes identified during sequencing of the blocks

Designation	Origin/nature of the nucleotide change
M	Putative mutation arising during PCR
PE	Generated during PCR as a consequence of an error in the sequence of a primer
SE	Putative error in the published sequence of the gene
SM	Silent mutation engineered into relevant blocks during PCR

5.3.6a Sequencing amyLQS50 *and wild type* amyL *block 1 plasmids*

The plasmids from four LQS1 clones and two WT1 clones were sequenced (Table 5.8). The DNA sequences, along with the nucleotides identified during this study being at variance with the published sequence are shown in Fig. 5.13. Plasmid pLQS1-2 differed from the expected sequence at four nucleotides (SE1, M1, M2 and M3), three of which resulted in alterations in the encoded amino acid (Table 5.10). Likewise, the determined sequence of pLQS1-5 differed from the expected sequence at three nucleotides (SE1, M2, M4) which also lead to changes in the corresponding amino acids of the proteins. The least number of nucleotide alterations were found in pLQS1-7 which had only one nucleotide change (SE1) when compared with the expected sequence. Finally, the *amyL* block 1 plasmid clones pWT1-11 and pWT1-12 deviated from the published sequence of the *amyL* gene at only one nucleotide (SE1) which also changed the amino acid sequence of the protein (Table 5.11).

When compared to the expected sequence of *amyLQS50* and that published for *amyL* (Gray *et al.*, 1986), all of the PCR generated block 1 fragments had one nucleotide change in common, SE1. This was not located in a region to which a PCR primer annealed and therefore, could not have been generated by a primer sequence error. It is likely that this nucleotide alteration was not introduced during PCR but represented an error in the published sequence of the *amyL* gene. The M2 nucleotide change was identified in pLQS1-2 and pLQS1-5 but not in pLQS1-7, suggesting that this was an error generated during PCR.

Sequencing of the LQS1 clones confirmed that the SOE reactions had proceeded as planned, resulting in the replacement of a region of the *amyL* gene with that of *amyS*. Discrepancies with the expected sequence were found in all *amyLQS50* clones. However, since nucleotide change SE1 appeared to result from an error in the sequence of *amyL*, and not an error generated during PCR, pLQS1-7 was chosen as the clone from which to isolate the LQS1 fragment for

the gene assembly process. Both of the WT1 plasmids had only one nucleotide change, SE1, and the plasmid chosen for gene assembly was pWT1-11. The sequencing data also confirmed that the *Eco*RI and *Sac*I restriction sites had been created in the ends of the PCR products.

Continued

Block 1

amyLQS50

Eco*RI** **Pst*I

GAATTCTTGCTGCCTCATT**CTGCAG**CAGCGGCGGCAAATC
TTAATGGGACGCTGATGCAGTATTTTGAATGGTACATGCC
CAATGACGGCCAACATTGGA(M1)A(SE1)GCGTTTGCAAAACGA(M4)
CTCGGCATATTTGGCTGAACACGGTATTACTGCCGTCTGGA
TTCCCCCGGCATATAAGGGAACAAGCCGCAGCGACGTAGG
GTACG(M2)GTGCTTACGACCTTTATGATTTAGGGGAGTTTCAT
CAAAAAGGGACGGTTCGGACAAAGTACGGCACAAAAGGAG
AGCTGCAATCTGCGATCAAAAGTCTTCAT(M3)TCCCGCGACAT
TAACGTTTACGGGGATGTGGTCATCAACCACAAAGGCGGCG
CTGATGCGACCGAAGATGTAA**CCGCGG**TTGAA**GAGCTC**
***Sac*II** ***Sac*I**

amyL

Eco*RI** **Pst*I

GAATTCTTGCTGCCTCATT**CTGCAG**CAGCGGCGGCAAATC
TTAATGGGACGCTGATGCAGTATTTTGAATGGTACATGCC
CAATGACGGCCAACATTGGAA(SE1)GCGTTTGCAAAACGACT
CGGCATATTTGGCTGAACACGGTATTACTGCCGTCTGGATT
CCCCCGGCATATAAGGGAACGAGCCAAGCGGATGTGGGCT
ACGGTGCTTACGACCTTTATGATTTAGGGGAGTTTCATCAA
AAAGGGACGGTTCGGACAAAGTACGGCACAAAAGGAGAG
CTGCAATCTGCGATCAAAAGTCTTCATTCCCGCGACATTAA
CGTTTACGGGGATGTGGTCATCAACCACAAAGGCGGCGCT
GATGCGACCGAAGATGTAA**CCGCGG**TTGAA**GAGCTC**
***Sac*II** ***Sac*I**

Fig. 5.13. The expected DNA sequences of amy*LQS50* and *amyL* block 1 fragments.

Restriction sites are indicated in bold face type and sites incorporated during PCR are underlined. The positions of the nucleotide changes identified during sequencing are shown in red.

Table 5.10. The nucleotide changes in the *amyLQS50* block 1 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino acid change
pLQS1-2	A → G	M1	AAG → GAG	Lys → Glu
	A → G	SE1	AAG → AGG	Lys → Arg
	G → A	M2	GGT → AGT	Gly → Ser
	T → C	M3	CAT → CAC	His → His
pLQS1-5	A → G	SE1	AAG → AGG	Lys → Arg
	G → A	M2	GGT → AGT	Gly → Ser
	A → G	M4	GAC → GGC	Asp → Gly
pLQS1-7	A → G	SE1	AAG → AGG	Lys → Arg

Table 5.11. The nucleotide changes in the *amyL* block 1 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino Acid change
pWT1-11	A → G	SE1	AAG → AGG	Lys → Arg
pWT1-12	A → G	SE1	AAG → AGG	Lys → Arg

5.3.6b Sequencing amyLQS50 *and wild type* amyL *block 2 plasmids*

Two LQS2 plasmids and two WT2 plasmids (Table 5.8) were sequenced during this study. With the exception of pLQS2-3, the sequences of the plasmids were as expected (Fig. 5.14). The DNA fragment in pLQS2-3 exhibited a single base pair deletion (DEL1, Fig. 5.14) however, the alternative LQS2 plasmid, pLQS2-4, did not contain this deletion. The DNA sequences obtained from pWT2-8 and pWT2-9 were identical to the published sequence of the *amyL* gene. Additionally, the sequence data confirmed the incorporation of *Sac*I and *Bam*HI restriction sites into the ends of the fragments during PCR. The plasmids selected to provide the block 2 DNA fragment for gene assembly were pLQS2-4 and pWT2-9.

Block 2

amyLQS50

***Sac*I** ***Sac*II**
GAGCTCTGTAACCGCGGTTGAAGTCGATCCCGCTGACCGC
AACCGCGTAATTTCAGGAGAACACCTAATTAAAGCCTGGA
CACATTTTCATTTTCCCGGGCGGGGCAACACCTACTCC(DEL1)A
GCTTTAAGTGGCGCTGGTACCATTTTGACGGAACCGATTGG
GACGAGTCCCGAAAGCTGAACCGCATCTATAAGTTTCAAGG
AAAGGCTTGGGATTGGGAAGTTTCCAATGAAAACGGCAACT
ATGATTATTTGATGTATGCCGAC**ATCGAT**TATGACCA**GGATCC**
***Cla*I** ***Bam*HI**

amyL

***Sac*I** ***Sac*II**
GAGCTCTGTAACCGCGGTTGAAGTCGATCCCGCTGACCGC
AACCGCGTAATTTCAGGAGAACACCTAATTAAAGCCTGGA
CACATTTTCATTTTCCGGGGCGCGGCAGCACATACAGCGAT
TTTAAATGGCATTGGTACCATTTTGACGGAACCGATTGGGA
CGAGTCCCGAAAGCTGAACCGCATCTATAAGTTTCAAGGA
AAGGCTTGGGATTGGGAAGTTTCCAATGAAAACGGCAACT
ATGATTATTTGATGTATGCCGAC**ATCGAT**TATGACCA**GGATCC**
***Cla*I** ***Bam*HI**

Fig. 5.14. The expected DNA sequences of amy*LQS50* and *amyL* block 2 fragments.

Restriction sites are indicated in bold face type and sites incorporated during PCR are underlined. The deleted nucleotide is indicated in red.

5.3.6c Sequencing **amyLQS50** *and wild type* **amyL** *block 3 plasmids*

To obtain suitable clones for assembly of the amylase genes two LQS3 plasmids and two WT3 plasmids were sequenced (Table 5.8). Deviations from the expected sequence (Fig. 5.15) are shown in Table 5.12.

pLQS3-5 exhibited three nucleotide changes (M6, M7 and M16) when compared with the expected sequence. M6 and M16 altered the encoded amino acid (Table 5.12). In contrast, M7 was a conservative change and had no affect on the encoded amino acid. pLQS3-6 was found to differ from the expected sequence only at nucleotide position, M16 (Fig. 5.15 and Table 5.12). The equivalent nucleotide change was also detected in pLQS3-5 and it is possible that this represented an error in the published sequence of the *amyS* gene. However, this could not be verified since too few independent LQS3 clones were sequenced.

The DNA sequence of LQS3 DNA fragments contained within pLQS3-5 and pLQS3-6 confirmed that the correct splicing reactions had occurred. In spite of the nucleotide change observed in this clone (M16), pLQS3-6 was selected to provide the *amyLQS50* block 3 fragment for subsequent gene assemblies.

Sequencing of the WT3 fragments contained within pWT3-15 and pWT3-16 did not reveal any deviations from the published sequence of the *amyL* gene and the former plasmid was selected for gene assembly. Finally, the sequence data obtained for all the block 3 fragments verified the presence of the *Bam*HI restriction site which was incorporated during PCR.

Continued

Block 3

amyLQS50

***Bam*HI** ***Cla*I**

GGATCCATGCCGAC**ATCGAT**TATGACCATCCTGATGTCGC
AGCAGAAATTAAGAGATGGGGCAC TTGGTATGCCAATGAA
CTGCAATTGGACGGTTTCCGTCTTGATGCTGTCAAACACATT
AAATTTTCTTTTTTGCGGGATTGGG(M6)TTAATCATGTCAGGGA
AAAAACTGGCAAGCCGCTATTTACCGTCGGGGAATATTGGA
GCTATGACATCAACAAGT(M7)TGCACAATTACATTACGAAAAC
AA(M16)GCGGAACGATGTCTTTGTTTGATGCCCCGTTACACAACA
AATTTTATACCGCTTCCAAATCAGGAGGCGGCTATGATATGAG
GAAATTGCTGAACGGTACGGTCGTTTCCAAGCATCCGTTGAAA
TCGGTTACATTTGTCGATAACCATGATACACAGCCGGGGCAA
TCGCTTGAGTCGAC**GTCGAC**

***Sal*I**

amyL

***Bam*HI** ***Cla*I**

GGATCCATGCCGAC**ATCGAT**TATGACCATCCTGATGTCGC
AGCAGAAATTAAGAGATGGGGCACTTGGTATGCCAATGAA
CTGCAATTGGACGGTTTCCGTCTTGATGCTGTCAAACACAT
TAAATTTTCTTTTTTGCGGGATTGGGTTAATCATGTCAGGG
AAAAAACGGGGAAGGAAATGTTTACGGTAGCTGAATATTG
GCAGAATGACTTGGGCGCGCTGGAAAACTATTTGAACAAA
ACAAATTTTAATCATTCAGTGTTTGACGTGCCGCTTCATTAT
CAGTTCCATGCTGCATCGACACAGGGAGGCGGCTATGATAT
GAGGAAATTGCTGAACGGTACGGTCGTTTCCAAGCATCCGT
TGAAATCGGTTACATTTGTCGATAACCATGATACACAGCCG
GGGCAATCGCTTGAGTCGAC**GTCGAC**

***Sal*I**

Fig. 5.15. The expected DNA sequences of amy*LQS50* and *amyL* block 3 fragments.

Restriction sites are indicated in bold face type and sites incorporated during PCR are underlined. The positions of the nucleotide changes identified during sequencing are shown in red.

Table 5.12. The nucleotide changes in the *amyLQS50* block 3 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino acid change
pLQS3-5	G → C	M6	GTT → CTT	Val → Leu
	T → C	M7	TTG → CTG	Leu → Leu
	A → G	M16	AAC → GAC	Asn → Asp
pLQS3-6	A → G	M16	AAC → GAC	Asn → Asp

5.3.6d Sequencing amyLQS50 *and wild type* amyL *block 4 plasmids*

Three LQS4 plasmids and two WT4 plasmids were sequenced during this study (Table 5.8). To facilitate the assembly of block 4 into the full length gene, a *Mlu*I site was incorporated into LQS4 and WT4, using a single silent mutation created during PCR. This was achieved by a T to C nucleotide change (SM1) which did not affect the encoded amino acid (Tables 5.13 and 5.14, Fig. 5.16). The designed SM1 nucleotide change was observed in both LQS4 and WT4 fragments. Sequencing of LQS4 and WT4 fragments also revealed a second nucleotide change (PE1; Tables 5.13 and 5.14, Fig. 5.16). Although PE1 was initially thought to be an error in the published sequence of *amyL*, closer inspection of primer-LQS50/18 revealed an error in its sequence which resulted in the incorporation of a nucleotide during PCR which was not complementary to the template. Fortuitously, this G to C nucleotide change did not alter the amino acid encoded by the respective codon.

Sequencing of pLQS4-7 and pLQS4-10 revealed, in addition to SM1 and PE1, another nucleotide change (M8) common to both plasmids which affected the amino acid codon (Table 5.13). The equivalent nucleotide change was not present in pLQS4-5 and therefore, M8 must have arisen as a result of a nucleotide misincorporation during PCR. As a consequence, pLQS4-5 was selected to provide the LQS4 fragment for the gene assembly process. Apart from PE1 (Table 5.14), the sequence of the WT4 plasmids were as expected (Fig. 5.16) and pWT4-18 was used for gene assembly. Finally, sequencing of all block 4 fragments confirmed the presence of the engineered *Sph*I recognition site which was incorporated during PCR.

Continued

Block 4

amyLQS50

***Sal*I**
GTCGACTGTCCAAACATGGTTTAAGCCGCTTGCTTACGCT
TTTATTCTCACAAGGGAATCTGGATACCCTCAGGTTTTCT
ACGGGGATATGTACGGGACGAAAGGGACATCGCCAAAG
GAAA(M8)TTCCCTCACTGAAACACAAAATTGAACCGATCTT
AAAAGCG(PE1)AGAAAACAGTAT(SM1)GCGTACGGAGCACAG
GCATGC
***Sph*I**

amyL

***Sal*I**
GTCGACTGTCCAAACATGGTTTAAGCCGCTTGCTTACGCT
TTTATTCTCACAAGGGAATCTGGATACCCTCAGGTTTTCT
ACGGGGATATGTACGGGACGAAAGGAGACTCCCAGCGCG
AAATTCCTGCCTTGAAACACAAAATTGAACCGATCTTAAA
AGCG(PE1)AGAAAACAGTAT(SM1)GCGTACGGAGCACAG**GCATGC**
***Sph*I**

Fig. 5.16. The expected DNA sequences of amy*LQS50* and *amyL* block 4 fragments.

Restriction sites are indicated in bold face type and sites incorporated during PCR are underlined. The positions of the nucleotide changes identified during sequencing are shown in red.

Table 5.13. The nucleotide changes in the *amyLQS50* block 4 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino acid change
pLQS4-5	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → C	PE1	GCG → GCC	Ala → Ala
pLQS4-7	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → C	PE1	GCG → GCC	Ala → Ala
	A → T	M8	GAA → GAT	Glu → Asp
pLQS4-10	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → C	PE1	GCG → GCC	Ala → Ala
	A → T	M8	GAA → GAT	Glu → Asp

Table 5.14. The nucleotide changes in the *amyL* block 4 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino acid change
pWT4-18	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → C	PE1	GCG → GCC	Ala → Ala
pWT5-20	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → C	PE1	GCG → GCC	Ala → Ala

5.3.6e Sequencing amyLQS50 *and wild type* amyL *block 5 plasmids*

To identify block 5 clones suitable for use in the assembly process, four LQS5 plasmid and three WT5 plasmids were sequenced (Table 5.8). As for block 4, all the LQS5 and WT5 fragments possessed SM1, creating a *Mlu*I recognition site (Fig. 5.17 A, B, Tables 5.15 and 5.16). The four LQS5 plasmids contained another common nucleotide change, SE17, which did not affect the encoded amino acid. This could not be attributed to a primer error and is likely to represents an error in the published sequence of *amyS*.

pLQS5-25 had, in addition to SM1 and SE17, a third nucleotide change (M9) which altered the amino acid codon (Table 5.15). The block 5 fragment encoded by pLQS5-33 also had a third nucleotide change (M10) which resulted

in an amino acid change. The most nucleotide changes were found in pLQS5-39 which had four changes excluding SM1 and SE17, all of which altered the encoded amino acids. The only nucleotide changes identified in pLQS5-8 were SM1 and SE17 and therefore, this plasmid was chosen for use in gene assembly. Finally, sequencing of the LQS5 plasmids confirmed that the splicing steps had proceeded as planned resulting in the replacement of a region of the *amyL* gene with the corresponding, but significantly larger, region of *amyS*.

The three WT5 plasmids sequenced contained SM1 (Table 5.16) and, apart from pWT5-17, the sequences were as expected (Fig. 5.17 B). This plasmid contained an additional nucleotide change, M15, which did not affect the encoded amino acid (Table 5.16). pWT5-11 was chosen to provide the WT5 fragment for the gene assembly process. Sequencing of all the block 5 plasmids verified the presence of the *Bam*HI recognition site which was incorporated during PCR.

Continued

Block 5

amyLQS50

***Bam*HI**

<u>**GGATCC**</u>GCAGAAAACAGTAT(SM1)GCGTACGGAGCACAGCATGA
TTATTTCGACCACCATGACATTGTCGGC TGGACAAGGGAAGGC
GACAGCTCGGTTG(M12)CAAATTCAGGTTTGGCGGCATTAATAAC
AGACGGACCCG(M11)G(M10)TGGGGCAAAGCG AATGTACGTTGGCA
AACAACACGCC(SE17)GGAAAAGTGTTCTATGACCTTACCGGCAAC
CGGAGTGACACCGTCACCATCAACA GTGATGGATGGGGGGAA
TTCAAAGTCAATGGCGGTTCGGTTTCGGTTTGGGT(M13)TCCTAGA
AAAACGACCGTTTCTACCATC GCTCGGCCGAT(M9)CACAACCCGA
CCGTGGACTGGTGAATTCGTCCGTTGGACCGAACCACGGTTGGT
GGCATGGCCTT(M14)AGAAGAGCAGAGAGGACGGATTTCCTGAAG
GAAATCCGTTTTTTTATTTTGCCCGTCTTATAAATTTCTTTGATT
ACATTTTATAATTAATTTTAACAAAGTGTCATCAGCCCTCAGGA
AGGACTTGCTGACAGTTTGAATCGCATAGGTAAGGCGGGGATG
AAATGGCAACGTTATCTGATGTAGCAAAGAAAGCAAATGTGTC
GAAAATGACGGTATCGCGGGTGATCAATCATCCTGAGAC TGTG
ACGGATGAATTGAAAA**AAGCTT**

***Hind*III**

Fig. 5.17 A. The expected DNA sequence of the *amyLQS50* block 5 fragment.

Restriction sites are indicated in bold face type and sites incorporated during PCR are underlined. The positions of the nucleotide changes identified during sequencing are shown in red.

Block 5

amyL

***Bam*HI**

GGATCCGCAGAAAACAGTAT(SM1)GCGTACGGAGCACAGCATGA
TTATTTCGACCACCATGACATTGTCGGCTGGACAAGGGAAGGC
GACAGCTCGGTTGCAAATTCAGGTTTGGCGGCATTAATAACAG
ACGGACCCGGTGGGGCAAAGCGAATGTATGTCGGCCGGCAAA
ACGCCGGTGAGACATGGCATGACATTACCGGAAACCGTTCGGA
GCCGGTTGTCATCAATTCGGAAGGCTGGGGAGAGTTTCACGTA
AACGGCGGG(M15)TCGGTTTCAATTTATGTTCAAAGATAGAAGAG
CAGAGAGGACGGATTTCCTGAAGGAAATCCGTTTTTTTATTTTG
CCCGTCTTATAAATTTCTTTGATTACATTTTATAATTAATTTTAA
CAAAGTGTCATCAGCCCTCAGGAAGGACTTGCTGACAGTTTGA
ATCGCATAGGTAAGGCGGGGATGAAATGGCAACGTTATCTGAT
GTAGCAAAGAAAGCAAATGTGTCGAAAATGACGGTATCGCGG
GTGATCAATCATCCTGAGACTGTGACGGATGAATTGAAAA**AGCTT**

***Hind*III**

Fig. 5.17 B. The expected DNA sequence of the *amyL* block 5 fragment.

Restriction sites are indicated in bold face type and sites incorporated during PCR are underlined. The positions of the nucleotide changes identified during sequencing are shown in red.

Table 5.15. The nucleotide changes in the *amyLQS50* block 5 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino acid change
pLQS5-8	T → C	SM1	TAT → TAC	Tyr → Tyr
	C → T	SE17	GCC → GCT	Ala → Ala
pLQS5-25	T → C	SM1	TAT → TAC	Tyr → Tyr
	T → C	M9	ATC → ACC	Ile → Thr
	C → T	SE17	GCC → GCT	Ala → Ala
pLQS5-33	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → A	M10	GGT → GAT	Gly → Asp
	C → T	SE17	GCC → GCT	Ala → Ala
pLQS5-39	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → C	M11	GGT → CGT	Gly → Arg
	G → C	M12	GCA → CCA	Ala → Pro
	T → C	M13	GTT → GCT	Val → Ala
	T → A	M14	TAG → AAG	Stop → Lys
	C → T	SE17	GCC → GCT	Ala → Ala

Table 5.16. The nucleotide changes in the *amyL* block 5 clones

Plasmid	Nucleotide change	Designation	Codon change	Amino acid change
pWT5-11	T → C	SM1	TAT → TAC	Tyr → Tyr
pWT5-17	T → C	SM1	TAT → TAC	Tyr → Tyr
	G → A	M15	GGG → GGA	Gly → Gly
pWT5-18	T → C	SM1	TAT → TAC	Tyr → Tyr

5.3.6f Sequencing summary

Sequencing was carried out on plasmids containing the PCR-generated DNA fragments corresponding to the component *amyLQS50* and wild type *amyL*

blocks. Sequencing was used in the first instance to verify that the splicing reactions had taken place as planned during the construction of the *amyLQS50* blocks. Secondly, it was used to confirm the presence of the restriction sites which were incorporated during PCR. Finally, since *Taq* DNA polymerase has no proof-reading function, the sequencing was used to identify mutations in the blocks created by nucleotide misincorporation during PCR.

Sequencing of the plasmids revealed a significant number of unexpected nucleotide changes in the fragments encoding the blocks. Excluding engineered nucleotide changes and apparent errors in the original sequences, all but one of the nucleotide changes (M15) occurred in fragments generated by SOE. This was not unexpected since the SOE process involved multiple rounds of splicing and amplification, thereby increasing the probability of nucleotide misincorporation by *Taq*. The sequencing of several clones allowed PCR-generated mutations to be differentiated from putative errors in the published sequences. *amyL* and *amyLQS50* were assembled (5.3.7) from the gene fragments encoded by the plasmids shown in Table 5.17.

Table 5.17. Plasmids selected for use in gene assembly

	Plasmid	
	amyLQS50	*amyL*
Block 1	pLQS1-7	pWT1-11
Block 2	pLQS2-4	pWT2-9
Block 3	pLQS3-6	pWT3-15
Block 4	pLQS4-5	pWT4-18
Block 5	pLQS5-8	pWT5-11

5.3.7 Assembly of the *amyLQS50* and wild type *amyL* blocks

To assemble the complete *Pst*I to *Hind*III *amyLQS50* gene fragments, the individual *amyLQS50* blocks were ligated in the appropriate order (1 to 5) by inserting them stepwise into the MCS of pUC19. Similarly, the component *amyL* blocks were assembled to form a reconstructed wild type gene. An advantage of constructing the genes from several blocks was that individual *amyLQS50* blocks could be replaced by corresponding *amyL* blocks permitting the construction of a range of genes with different combinations of *amyLQS50* and *amyL* blocks.

5.3.7a The gene assembly strategy

The *Pst*I to *Hind*III gene fragments were assembled using the block 1 to 5 DNA inserts and encoded three C-terminal amino acids of the AmyL signal peptide and the mature regions of the chimeric or wild type α-amylases. A three-stage strategy was devised for the assembly of chimeric and wild type genes (Fig. 5.18), as discussed below. In all stages of gene assembly, the relevant plasmids (Table 5.17) were digested (3.5), fragments gel purified (3.7), ligated (3.8) and then used to transform chemically competent *E. coli* XL1-Blue (3.10.4). Recombinant plasmids were isolated from transformants (3.4.1) and the presence of the correct DNA inserts verified by restriction analysis (3.5).

Stage 1

In the first stage of the assembly process, block 2 was released by digestion with *Sac*II and *Bam*HI and then cloned into the equivalent restriction sites of a block 1-containing plasmid, resulting in the joining of blocks 1 and 2. Block 3 was then released using *Cla*I and *Sal*I and cloned into the newly constructed block 1:2 plasmid using the same restriction sites. The resultant plasmid, p1:2:3, encoded blocks 1, 2 and 3 assembled in the correct order in the MCS of pUC19.

The lack of appropriate restriction sites within pUC19 meant that it was not possible to insert blocks 4 and 5 in a step wise manner. However, the restriction sites engineered into the ends of the blocks during PCR allowed blocks 4 and 5 to be assembled in a separate step (stage 2) and then joined to form the complete gene (stage 3).

Stage 2

The second stage involved the assembly of blocks 4 and 5. A block 5 fragment was released with *Mlu*I and *Hind*III and cloned into a block 4-containing plasmid which had been digested with the same enzymes. The resulting plasmid, p4:5, was then used in stage 3 of the assembly process.

Stage 3

In the last stage in the assembly process, the fused blocks 4 and 5 were released as a *Sal*I and *Hind*III fragment and cloned into p1:2:3 from stage 1 which had been digested with the same restriction enzymes. The resulting recombinant plasmid, p1:2:3:4:5, contained all five blocks of the α-amylase gene joined together in the correct order. This fragment encoded the mature form of the α-amylase (chimeric or wild type) which could be released as a *Pst*I to *Hind*III fragment for cloning into an expression vector.

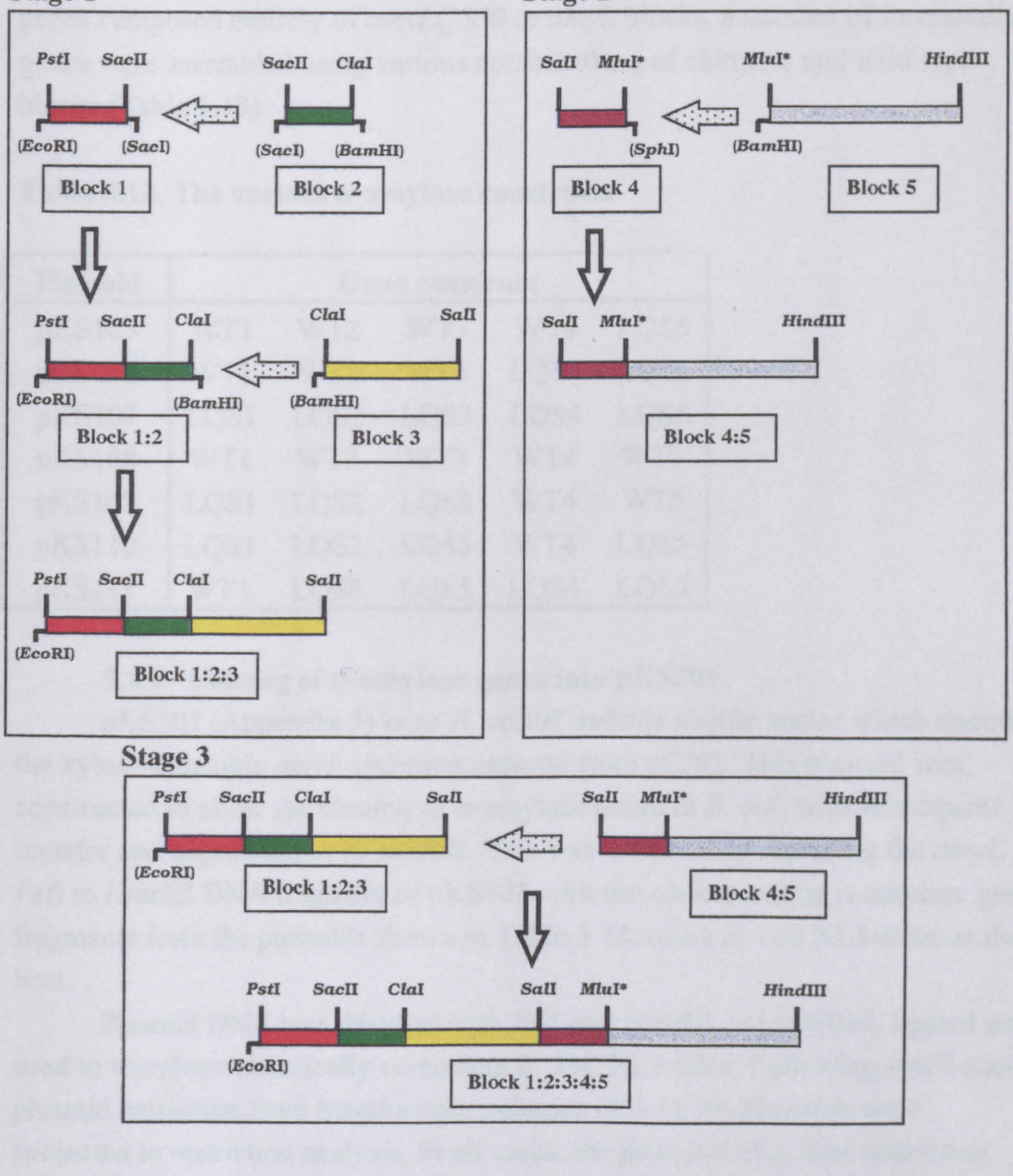


Fig. 5.18. The assembly of the five component blocks to construct the *Pst*I to *Hind*III α-amylase gene fragments.

The restriction sites above the gene/blocks are those within the *amyL* gene, except in the case of *Mlu*I* which was created by the introduction of a single silent mutation. Restriction sites below the blocks were those created during PCR to facilitate cloning and gene assembly.

5.3.7b Construction of genes with combinations of blocks

During the assembly process the five component *amyLQS50* blocks were combined in the correct order to form the complete chimeric gene. Similarly, the *amyL* blocks were assembled to reconstruct the wild type gene. In addition to the genes composed entirely of *amyLQS50* or *amyL* blocks, a number of intermediate genes were assembled using various combinations of chimeric and wild type blocks (Table 5.18).

Table 5.18. The various α-amylase constructs

Plasmid	Gene construct				
pKS105	WT1	WT2	WT3	WT4	LQS5
pKS106	WT1	WT2	WT3	LQS4	LQS5
pKS107	LQS1	LQS2	LQS3	LQS4	LQS5
pKS108	WT1	WT2	WT3	WT4	WT5
pKS109	LQS1	LQS2	LQS3	WT4	WT5
pKS110	LQS1	LQS2	LQS3	WT4	LQS5
pKS111	WT1	LQS2	LQS3	LQS4	LQS5

5.3.8 Cloning of α-amylase genes into pKS301

pKS301 (Appendix 5) is an *E. coli/B. subtilis* shuttle vector which encodes the xylose-inducible *amyL* secretion cassette from pCJ92. This plasmid was constructed to allow the cloning of α-amylase genes in *E. coli* with subsequent transfer and expression in *B. subtilis*. This was achieved by replacing the *amyL* *Pst*I to *Hind*III DNA fragment of pKS301 with the corresponding α-amylase gene fragments from the plasmids shown in Table 5.18, using *E. coli* XL1-Blue as the host.

Plasmid DNA was digested with *Pst*I and *Hind*III, gel purified, ligated and used to transform chemically competent *E. coli* XL1-Blue. Following small scale plasmid extraction from transformant colonies (3.4.1), the plasmids were subjected to restriction analysis. In all cases, the presence of a *Mlu*I restriction site within the cloned α-amylase genes was taken as evidence that the *amyL* gene of pKS301 had been successfully replaced. The α-amylase genes were expressed from plasmid- (6.1) and chromosome-based (7.4) expression systems and the genes were given the designations shown in Table 5.19.

Continued

Table 5.19. α-Amylase constructs encoded by shuttle and integrational plasmids

α-Amylase gene construct					Shuttle plasmid	Integrational plasmid	Gene designation
Native *amyL* gene					pKS301	pKS401	*amyL*
WT1	WT2	WT3	WT4	LQS5	pKS302	pKS402	*amyLQS50.1*
WT1	WT2	WT3	LQS4	LQS5	pKS303	-	*amyLQS50.2*
LQS1	LQS2	LQS3	LQS4	LQS5	pKS304	pKS404	*amyLQS50*
WT1	WT2	WT3	WT4	WT5	pKS305	pKS408	*amyL*
LQS1	LQS2	LQS3	WT4	WT5	pKS306	-	*amyLQS50.3*
LQS1	LQS2	LQS3	WT4	LQS5	pKS307	-	*amyLQS50.4*
WT1	LQS2	LQS3	LQS4	LQS5	pKS308	pKS405B	*amyLQS50.5*

CHAPTER 6

EXPRESSION OF CHIMERIC α-AMYLASES IN *BACILLUS SUBTILIS*

6.1 Transfer of the shuttle plasmids into *B. subtilis* DN1885

B. subtilis 168 and the majority of its derivatives encode an α-amylase, the product of the *amyE* gene. Although the native *B. subtilis* α-amylase is not closely related to those from *B. licheniformis*, *B. amyloliquefaciens* or *B. stearothermophilus*, the secretion of AmyL and the chimeric enzymes was studied in an α-amylase negative strain, *B. subtilis* DN1885 (Diderichsen *et al.*, 1990). The strain, provided by Novo Nordisk, was a derivative of *B. subtilis* RUB200 (Yoneda *et al.*, 1979).

Naturally competent *B. subtilis* DN1885 (3.9) was transformed with shuttle plasmids encoding chimeric and wild type α-amylases (5.3.8), selecting on plates containing chloramphenicol. The production of α-amylase by the DN1885 strains with plasmids encoding wild type or constructed α-amylase genes was investigated initially on penassay agar plates containing starch with or without 1% w/v xylose. After incubation at 37°C, zones of starch degradation were visualized by exposing the plates to iodine vapour (Fig. 6.1).

In the presence of xylose, zones of starch degradation were observed around strains containing plasmids pKS301, pKS302, pKS303, pKS305 and pKS308. Interestingly, in the absence of xylose, a large zone of starch hydrolysis was observed around colonies of DN1885 (pKS305), indicating that this plasmid had lost its xylose inducibility. Further, the induced starch hydrolysis associated with this strain was more extensive than that associated with the equivalent native *amyL* gene on pKS301 (Fig. 6.1). The production of α-amylase in the absence of xylose suggested that this plasmid had a mutation or had undergone a structural rearrangement during propagation. The secretion of active α-amylases by most of the strains confirmed the efficiency of the gene designs and splicing protocols.

On the basis of starch hydrolysis, α-amylase activity could not be detected in DN1885 (pKS304), DN1885 (pKS306) or DN1885 (pKS307). Therefore, whenever an α-amylase gene construct contained the LQS1 fragment (Table 5.19) a severe reduction in α-amylase activity was observed. It was not clear why it was not possible to detect the production of α-amylase from strains containing LQS1 since it could result from factors influencing any stage in production, from transcription to post-secretion release. It was also possible that the proteins produced lacked enzymatic activity. Since the starch plate assay is relatively insensitive and difficult to quantitate, α-amylase activity was measured directly in the supernatants of batch grown cultures (6.2).

- Xylose

+ Xylose

Fig. 6.1. Starch hydrolysis by *B. subtilis* DN1885 harbouring plasmids encoding wild type and chimeric α-amylases.

6.2 Production of α-amylase by batch grown cultures of *B. subtilis*

The plate assay provided preliminary information on the production of α-amylases by the various strains (6.1). In order to quantitate α-amylase production by the same strains, α-amylase activity in the supernatants of batch grown cultures was measured. 2xYT broth (11 ml) containing 1% w/v xylose and chloramphenicol was inoculated with resuspended bacterial colonies to an OD_{660} of 0.1. The cultures were incubated at 37°C with shaking (3.1) and at various time intervals samples were removed for the measurement of growth and α-amylase activity (3.14, Fig. 5.20).

The strains grew with a doubling time of approximately 42 min. In the case of DN1885 (pKS301), α-amylase was detected throughout exponential phase and into stationary phase. At t_0 the specific activity of α-amylase in the supernatant was 191 U/l/OD_{660} unit (Fig. 6.2 A). The α-amylase activity detected in the supernatant of this strain was used to define the wild type level against which the activity of the other strains was compared.

The activity of α-amylase detectable in the supernatant of DN1885 (pKS302) increased steadily throughout exponential phase and at t_0 the specific activity of α-amylase was 235 U/l/OD_{660} unit (Fig. 6.2 B). The specific activity of α-amylase continued to increase in stationary phase. α-Amylase activity could also be detected in the supernatant of DN1885 (pKS303) during exponential growth and at t_0 the specific activity was 51 U/l/OD_{660} unit.

α-Amylase activity was not detected in the supernatants of DN1885 (pKS304), DN1885 (pKS306) or DN1885 (pKS307) (Fig. 6.2 D, F and G). The failure to detect activity in these strains confirmed the results obtained with the plate assay (6.1). α-Amylase was detected in exponentially growing DN1885 (pKS305) and remained relatively constant throughout growth and increased slightly at t_0 to 521 U/l/OD_{660} unit (Fig. 6.2 E), a value higher than that obtained for the other strains studied. α-Amylase activity was not detected in the supernatant of DN1885 (pKS308) until the end of exponential growth and at t_0 the specific activity of α-amylase was 43 U/l/OD_{660} unit (Fig. 6.2 H).

Continued

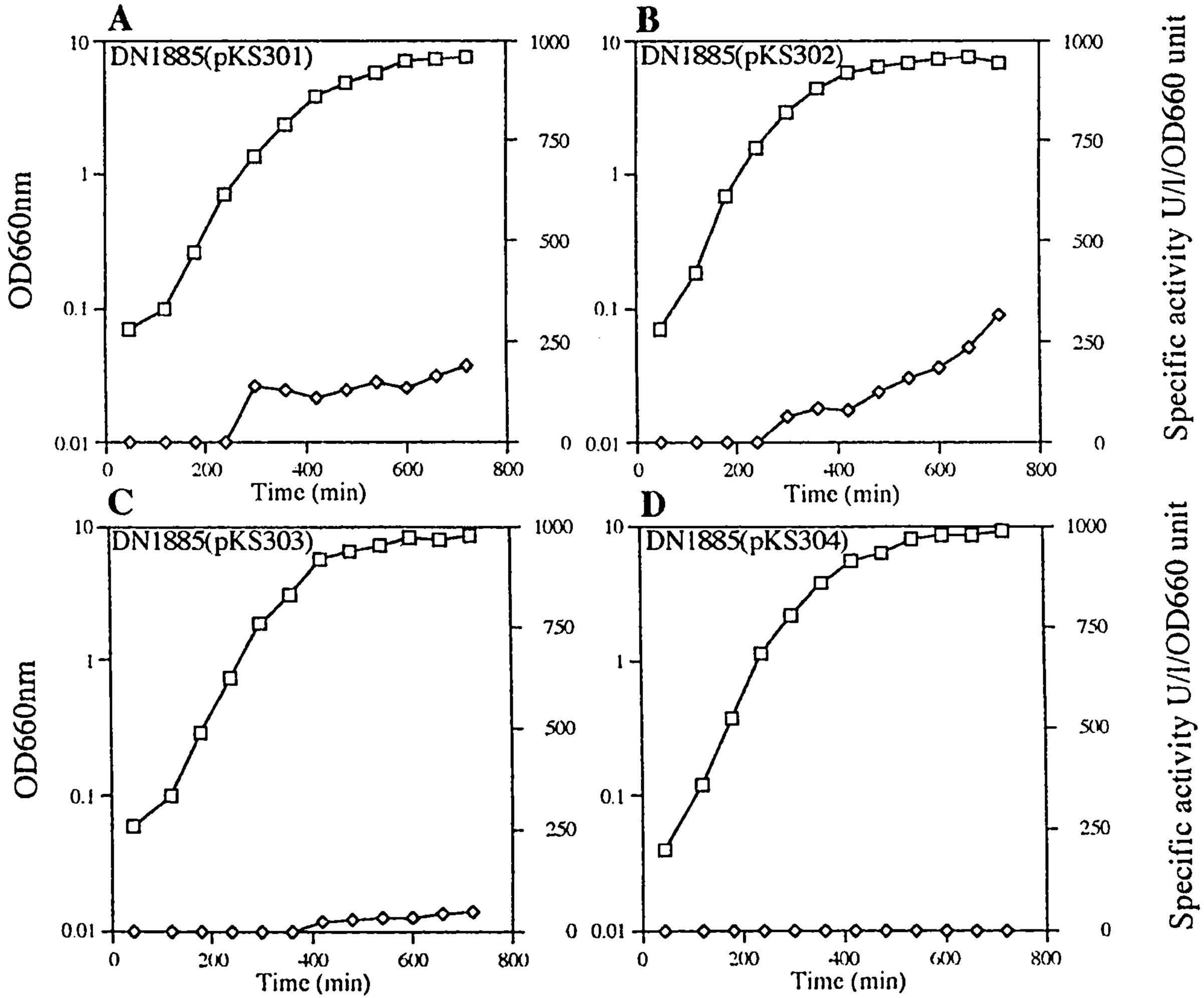


Fig. 6.2. Growth and specific activity of α-amylase in culture supernatants.
Key:
—□— Optical density
—◇— Specific activity

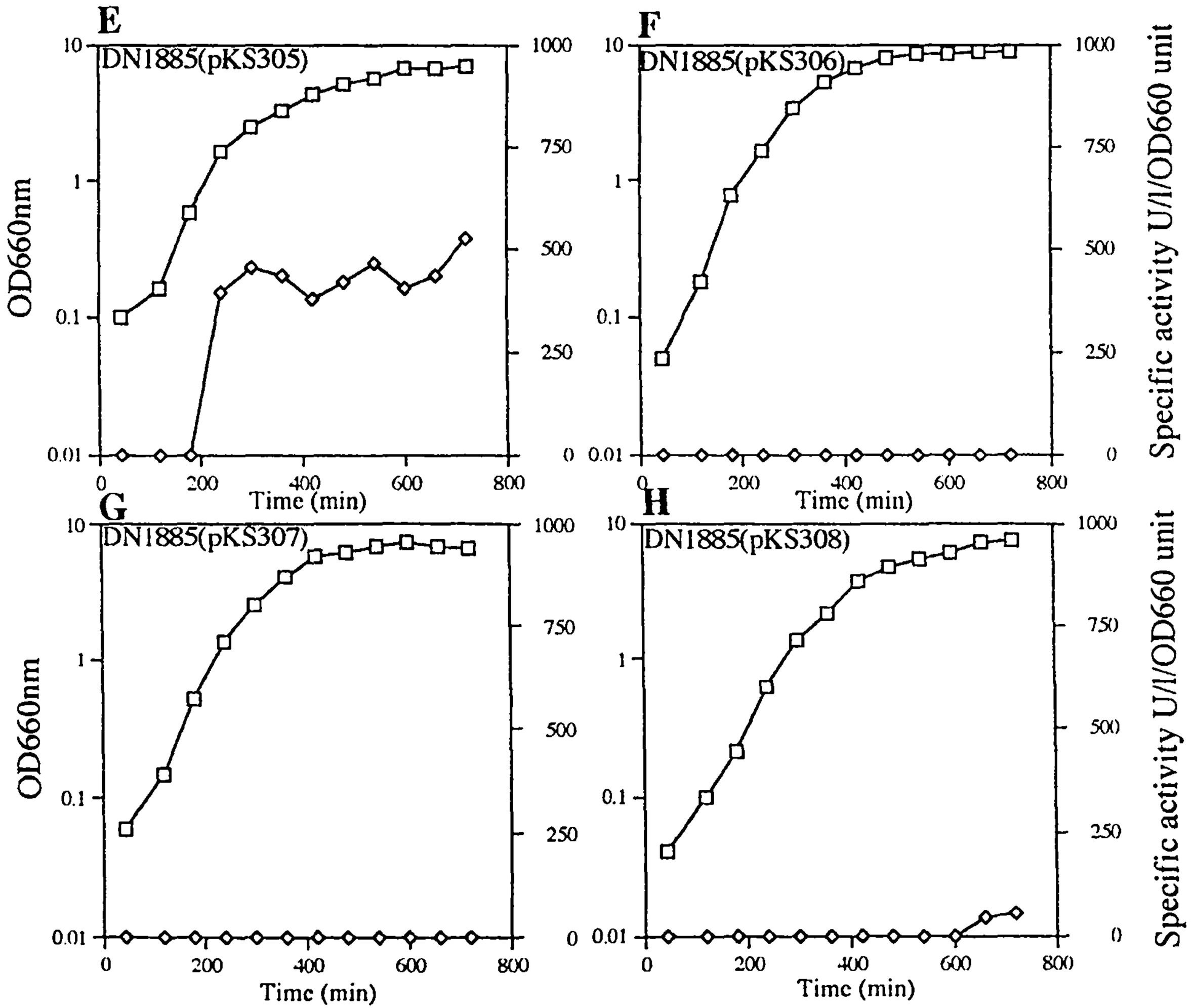


Fig. 6.2 continued. Growth and specific activity of α-amylase in culture supernatants.
Key:

—□— Optical density

—◇— Specific activity

The levels of α-amylase detected in the supernatants of strains harbouring plasmids varied considerably. This was not unexpected since, the α-amylase genes encoded by the plasmids were constructed using different combinations of *amyLQS50* and wild type *amyL* blocks. The levels of α-amylase activity detected in the supernatants of the different strains at t_0 in relation to DN1885 (pKS301) are summarized in Table 6.1.

Table 6.1. The relative specific activity of α-amylase in culture supernatants at t_0

Strain	Relative specific activity*
DN1885 (pKS301)	1.00
DN1885 (pKS302)	1.23
DN1885 (pKS303)	0.27
DN1885 (pKS304)	0.00
DN1885 (pKS305)	2.28
DN1885 (pKS306)	0.00
DN1885 (pKS307)	0.00
DN1885 (pKS308)	0.23

* Specific activity at t_0 relative to DN1885 (pKS301)

With the exception of the *B. subtilis* DN1885 strains harbouring plasmids pKS304, pKS306 or pKS307, α-amylase was detected in supernatants before stationary phase. This indicates that the α-amylase genes were not subjected to catabolite repression when grown in 2xYT.

When compared with DN1885 (pKS301), the levels of α-amylase activity detected in the supernatant of DN1885 (pKS302) at equivalent time points were almost identical. However, towards the end of exponential growth, the specific activity of α-amylase in the supernatant of the latter strain increased slightly so that 23% more α-amylase was detected in the supernatant of DN1885 (pKS302) at t_0. The chimeric α-amylase gene on pKS302 consisted of wild type *amyL* blocks apart from the LQS5 fragment (Table 5.19). These results show that the presence of the LQS5 gene fragment encoding the AmyS "tail" did not decrease the level of α-amylase activity detected in the supernatant of this strain.

The level of α-amylase detectable in the supernatant of DN1885 (pKS303) was much reduced when compared to the wild type strain DN1885 (pKS301) since, at t_0, the former strain exhibited only 27% of the activity of the latter. The

gene encoding AmyLQS50.2 was constructed from *amyL* blocks 1 to 3 in combination with *amyLQS50* blocks 4 and 5. This suggested that the presence of amino acid residues 372 to 381 from AmyQ (LQS4) and residues 438 to 515 from AmyS (LQS5) affected the secretion and/or activity of this enzyme.

pKS305 encoded the reconstructed *amyL* gene. As a consequence, the levels of α-amylase activity produced by DN1885 (pKS305) were expected to be similar to those detected in the supernatant of DN1885 (pKS301). However, this was not the case and at all phases of growth the amount of α-amylase produced by DN1885 (pKS305) was at least double that detected in the supernatant of DN1885 (pKS301) (Fig. 6.2 E and Table 6.1).

The sequence of the *Pst*I to *Hind*III *amyL* gene fragment encoded by plasmid pKS305 was determined during this study (5.3.6) and, apart from a putatively identified sequence error and two silent mutations, corresponded exactly to the published nucleotide sequence of this gene (Gray *et al.*, 1986). Therefore, the amino acid sequence of the α-amylase encoded by pKS305 was known to be the same as AmyL and deviations from this sequence could not be responsible for the greater than 2-fold increase in the amount of α-amylase detected in the supernatant of DN1885 (pKS305).

Although α-amylase activity was not determined in batch cultures of DN1885 (pKS305) grown without xylose, evidence from the plate assay (6.1) revealed that significant amounts of α-amylase was produced by this strain in the absence of the inducing sugar. It is possible that this effect was caused by structural rearrangements or mutation in the promoter/operator site or within the gene encoding the XylR repressor. The constitutive expression of the reconstructed *amyL* gene on pKS305 was not investigated further in the study. The lowest α-amylase activity was detected in the supernatant of DN1885 (pKS308) which exhibited only 23% of the activity of DN1885 (pKS301) at t_0 (Table 6.1).

In summary, using SOE a range of chimeric α-amylase genes were constructed and expressed in *B. subtilis* DN1885 using a shuttle plasmid system. The chimeric proteins were exported using the signal peptide from AmyL and the amount the active enzyme detectable in the culture supernatant in each case was shown to be variable. α-Amylase activity was not detected in all of the strains and, as a general rule, the relative activity of the α-amylases was inversely proportional to the number of *amyLQS50* block making up the genes. Whenever the LQS1 fragment was included in a gene construct no α-amylase activity was detected, suggesting that the inclusion of this block lead to the production of an inactive or non-secretable enzyme. The exact reasons for the lower levels of α-

amylase activity detectable in certain culture supernatants was not possible to determine at this point, but were the focus of subsequent studies.

6.3 Western blotting of α-amylases in culture supernatants

The relative amount of α-amylases present in culture supernatants was investigated by Western blotting. Batch cultures of *B. subtilis* DN1885 and DN1885 with plasmid-encoded α-amylase genes were grown in 2xYT with 1% w/v xylose and chloramphenicol (20 μg/ml) for 48 hours (3.1). The cells were removed by centrifugation and supernatants subjected to SDS-PAGE (3.15) and Western blotting (3.16). The primary antibody was a polyclonal antiserum raised against AmyL (anti-Termamyl). Since the chimeric α-amylases were composed of large regions of AmyL, it was assumed that they would cross-react sufficiently to allow their detection.

Bands corresponding to α-amylases were observed in the Termamyl AmyL standard and the supernatant samples from *B. subtilis* DN1885 strains harbouring plasmids pKS302, pKS303, pKS305, pKS306 and pKS308 (Fig. 6.3). However, the intensities of the bands in the supernatants of the different strains varied considerably. The calculated molecular masses of the α-amylases were as expected (Table 7.1) and in the case of the chimeric proteins, were larger than AmyL. An intense band corresponding to AmyLQS50.1 was detected in the supernatant of DN1885 (pKS302) and a much fainter AmyLQS50.2 band in the supernatant of DN1885 (pKS303). The most intense α-amylase band was detected in the supernatant of DN1885 (pKS305).

A very faint AmyLQS50.3 band was observed in the supernatant of DN1885 (pKS306) but no α-amylase activity was detected in batch cultures of this strain (6.2). AmyLQS50.5 was also detected in the supernatant of DN1885 (pKS308). As expected, an α-amylase band was not detected in the supernatant of *B. subtilis* DN1885, confirming that the bands observed in the supernatants of plasmid-containing strains were specific for α-amylases and not caused by cross-reacting host proteins. Finally, no α-amylase bands were detected in the supernatants of DN1885 (pKS304) or DN1885 (pKS307) (Fig. 6.3).

Continued

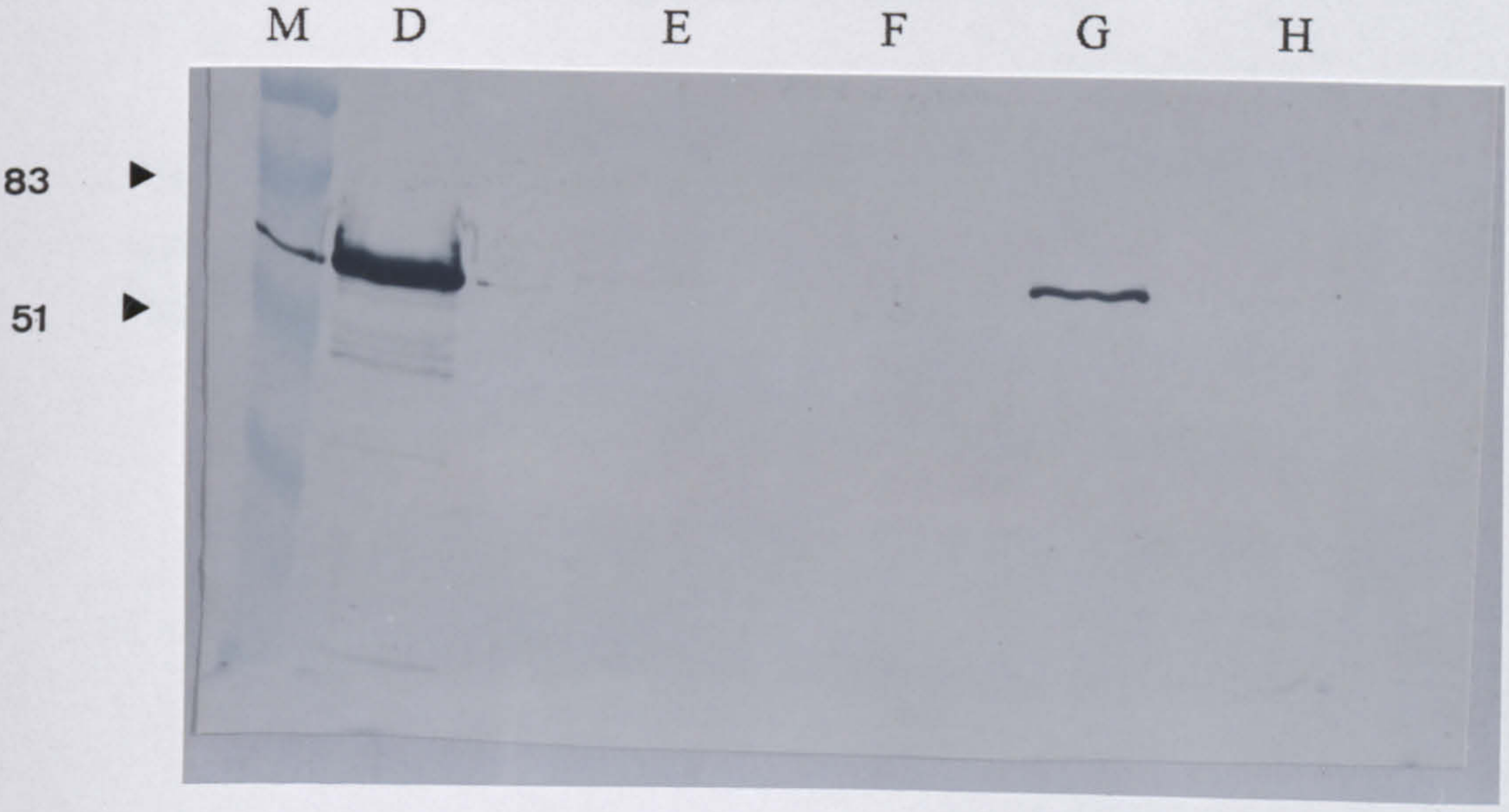


Fig. 6.3. Detection of α-amylases in culture supernatants by Western blotting.

Key:

M - Pre-stained markers (kDa)
T - Termamyl α-amylase standard
A - *B. subtilis* DN1885(pKS302)
B - *B. subtilis* DN1885(pKS303)
C - *B. subtilis* DN1885(pKS304)
D - *B. subtilis* DN1885(pKS305)
E - *B. subtilis* DN1885(pKS306)
F - *B. subtilis* DN1885(pKS307)
G - *B. subtilis* DN1885(pKS308)
H - *B. subtilis* DN1885

The anti-AmyL antiserum used in the Western blots does not necessarily cross-react with chimeric α-amylases to the same extent as AmyL and this factor alone may have contributed to the failure to detect α-amylases in the supernatants of some strains.

Notwithstanding the above, the relative amounts of cross-reacting protein detected in the culture supernatants correlated well with the α-amylase activity measured in the respective culture supernatants (6.2). Since AmyLQS50.5 produced an easily detectable band on the Western blots, it is unlikely that the failure to detect other chimeric α-amylases was the result of a lack of cross-reactivity with the primary antibody.

B. subtilis secretes numerous proteolytic enzymes at the end of exponential growth and in stationary phase (2.15). The presence of extracellular proteases can be a contributory factor in low product recovery when *B. subtilis* is used to export certain heterologous proteins (Wu *et al.*, 1991). Consequently, it is possible that the low levels of chimeric α-amylases detected in culture supernatants was a result of proteolytic degradation. The stability of AmyL and AmyLQS50.5 in culture supernatants was the focus of subsequent experiments (7.3).

The chimeric α-amylases present on pKS302 and pKS308, AmyLQS50.1 and AmyLQS50.5, respectively, could be detected in culture supernatants in significant quantities and therefore, these enzymes were selected for further study.

CHAPTER 7

PURIFICATION AND CHARACTERIZATION OF WILD TYPE AND CHIMERIC α-AMYLASES

7.1 Investigation of the pI of AmyLQS50.1, AmyLQS50.5 and AmyL

The calculated pI values and molecular masses of the chimeric proteins are shown in Table 7.1. Calculated pI values only provide an indication of a proteins charge and actual pI values must be determined by IEF or related techniques. IEF was used to investigate the overall charge of AmyLQS50.1, AmyLQS50.5 and AmyL.

Table 7.1. The calculated pI values and molecular masses for AmyL and chimeric α-amylases

α-Amylase	Calculated pI value*	Molecular mass (kDa)*
AmyL	6.3	55.11
AmyLQS50.1	6.5	58.85
AmyLQS50.2	6.6	58.78
AmyLQS50.3	6.9	55.21
AmyLQS50.4	7.8	58.95
AmyLQS50.5	7.8	58.84
AmyLQS50	8.1	58.88

* Calculated using the Lasergene software (DNASTAR Inc.)

α-Amylases in the supernatants of *B. subtilis* DN1885 (pKS302), DN1885 (pKS305) and DN1885 (pKS308) were analysed using IEF (pH range 3-10) and zymography (3.17). The relative positions of the α-amylases in the gel following IEF are indicated by zones of starch hydrolysis (Fig. 7.1). As expected, the α-amylase with the lowest pI was the wild type enzyme, AmyL, and that with the highest pI was AmyLQS50.5. AmyLQS50.1 had a pI which was intermediate between that of the other two α-amylases, demonstrating that the inclusion of the LQS5 fragment in the gene construct (Table 5.19) produced a marked change in the pI of the encoded α-amylase, without adversely affecting its activity (6.2).

This technique did not allow an accurate determination of the pI's of the α-amylases but it did confirm that the pI's of chimeric enzymes were increased with respect to the wild type. The exact pI's of the α-amylases were determined using the Rotofor apparatus (7.7).

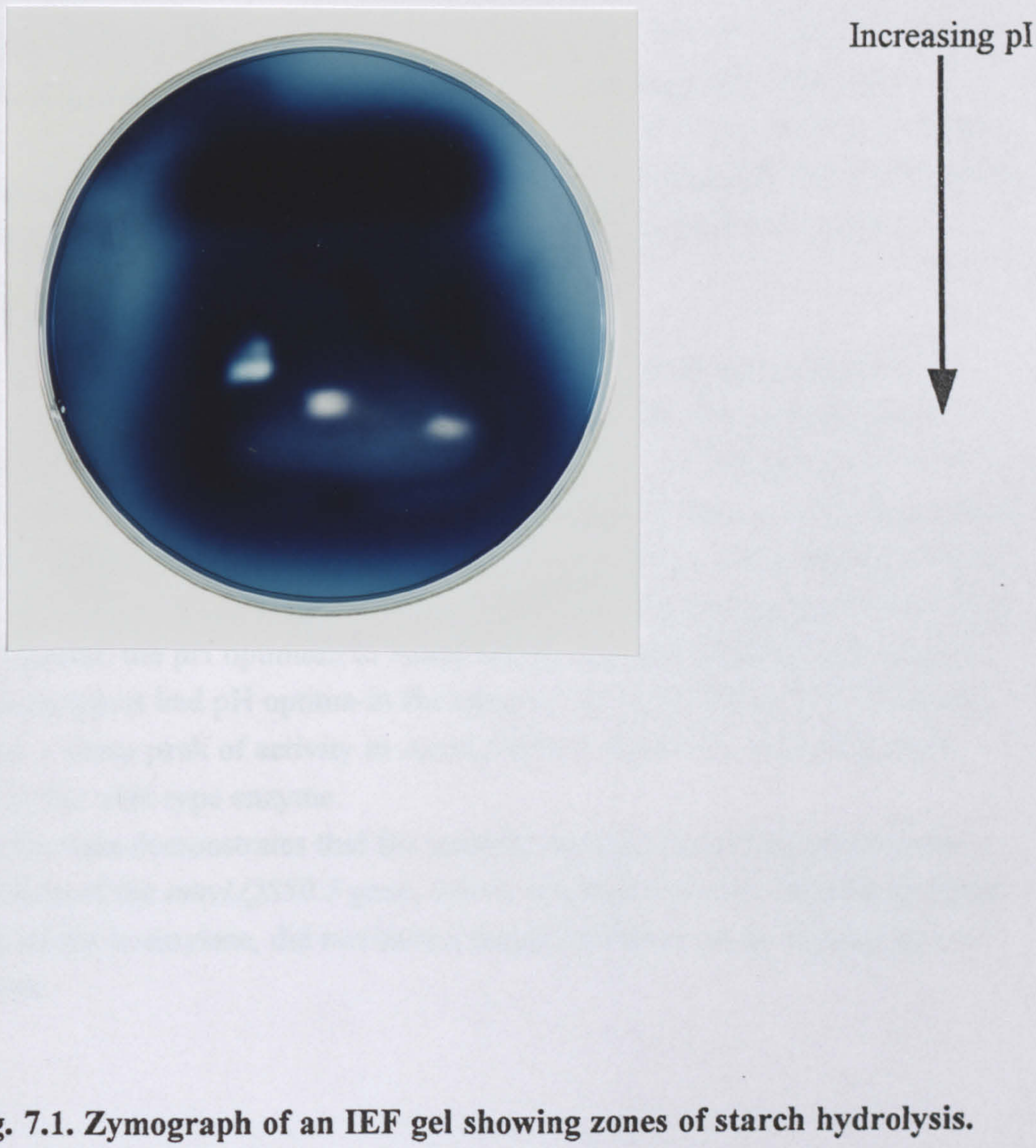


Fig. 7.1. Zymograph of an IEF gel showing zones of starch hydrolysis.
Key:
A - *B. subtilis* DN1885(pKS305)
B - *B. subtilis* DN1885(pKS302)
C - *B. subtilis* DN1885(pKS308)

7.2 Determination of the temperature optima, pH optima and thermostabilities of AmyLQS50.5 and AmyL

An important feature of the α-amylase from *B. licheniformis*, and one which makes this enzyme commercially important, is its ability to hydrolyze starch at temperatures in excess of 80°C (Suzuki *et al.*, 1989). The temperature optimum of this enzyme is in the range 90°C to 95°C and the enhanced thermostability of AmyL compared with the α-amylases from *B. amyloliquefaciens* and *B. stearothermophilus* has been demonstrated previously (Tomazic and Klibanov, 1988; Suzuki *et al.*, 1989).

Suzuki *et al* (1989) constructed chimeric α-amylases using the enzymes from *B. licheniformis* and *B. amyloliquefaciens* which did not exhibit the same temperature optima or thermostability profiles when compared with AmyL. Therefore, to characterize the properties of AmyLQS50.5 with respect to AmyL, the temperature optima (3.18), pH optima (3.19) and thermostability (3.20) of the enzymes in *B. subtilis* DN1885 (pKS305) and DN1885 (pKS308) culture supernatants were determined.

7.2.1 Temperature and pH optima of AmyL and AmyLQS50.5

When expressed as a function of temperature, the activity profiles of AmyL and AmyLQS50.5 were almost identical (Fig. 7.2). The activities of the enzymes in culture supernatants were found to increase steadily with temperature until the maximum values were reached at 90°C which then decreased markedly at 98°C. The temperature optimum of both enzymes was estimated at 90°C. In a similar manner, the pH optimum of AmyLQS50.5 was not altered significantly and both enzymes had pH optima in the range of 7.5 to 8.0 (Fig. 7.3). However, there was a sharp peak of activity in AmyLQS50.5 at pH 7.7, which was not evident in the wild type enzyme.

This data demonstrates that the specific changes introduced during the construction of the *amyLQS50.5* gene, which resulted in a considerable increase in the pI of the α-amylase, did not have a major influence on its temperature or pH optima.

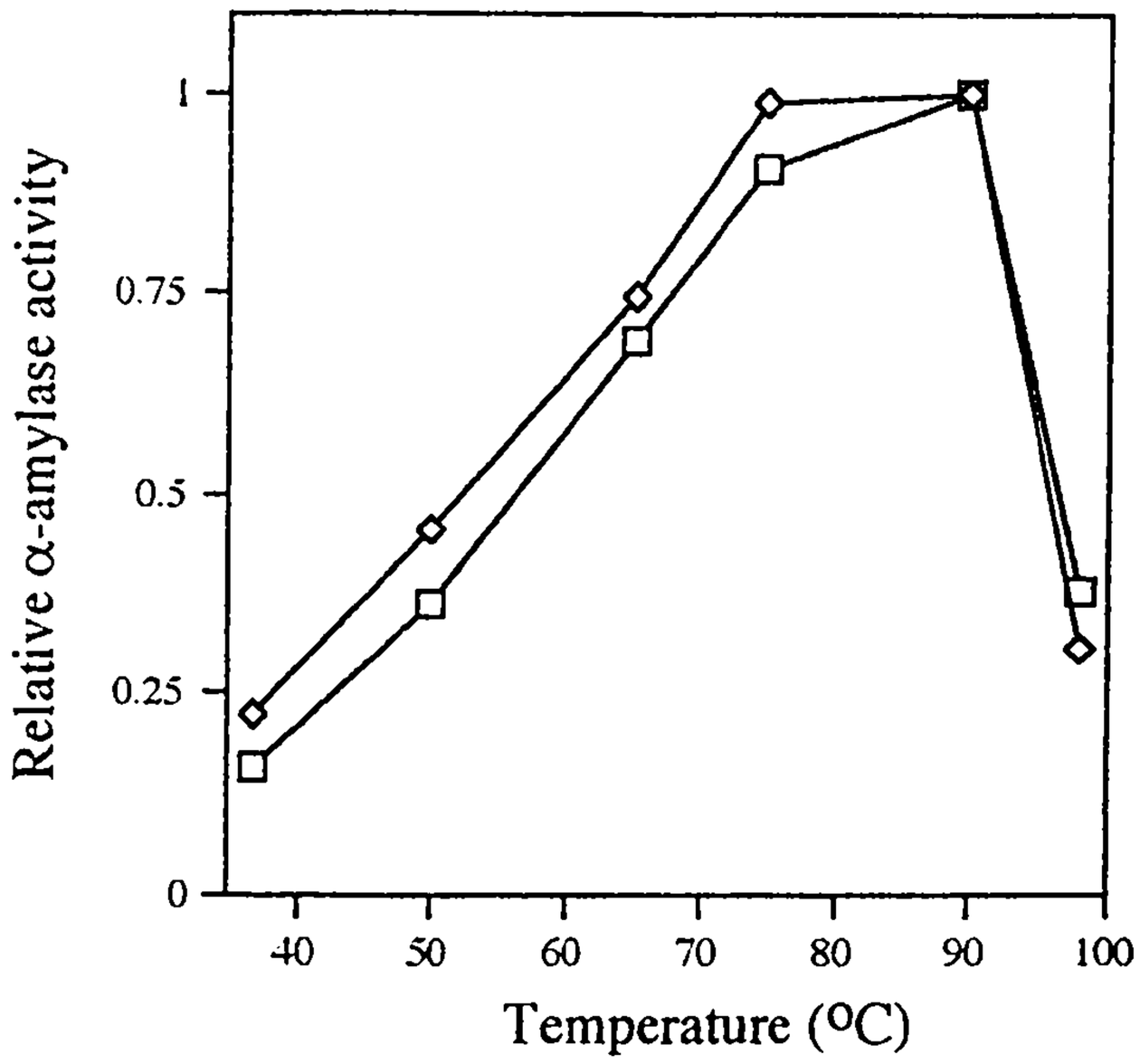


Fig. 7.2. Temperature optima of AmyLQS50.5 and AmyL.
Key:
—□— AmyL
—◇— AmyLQS50.5

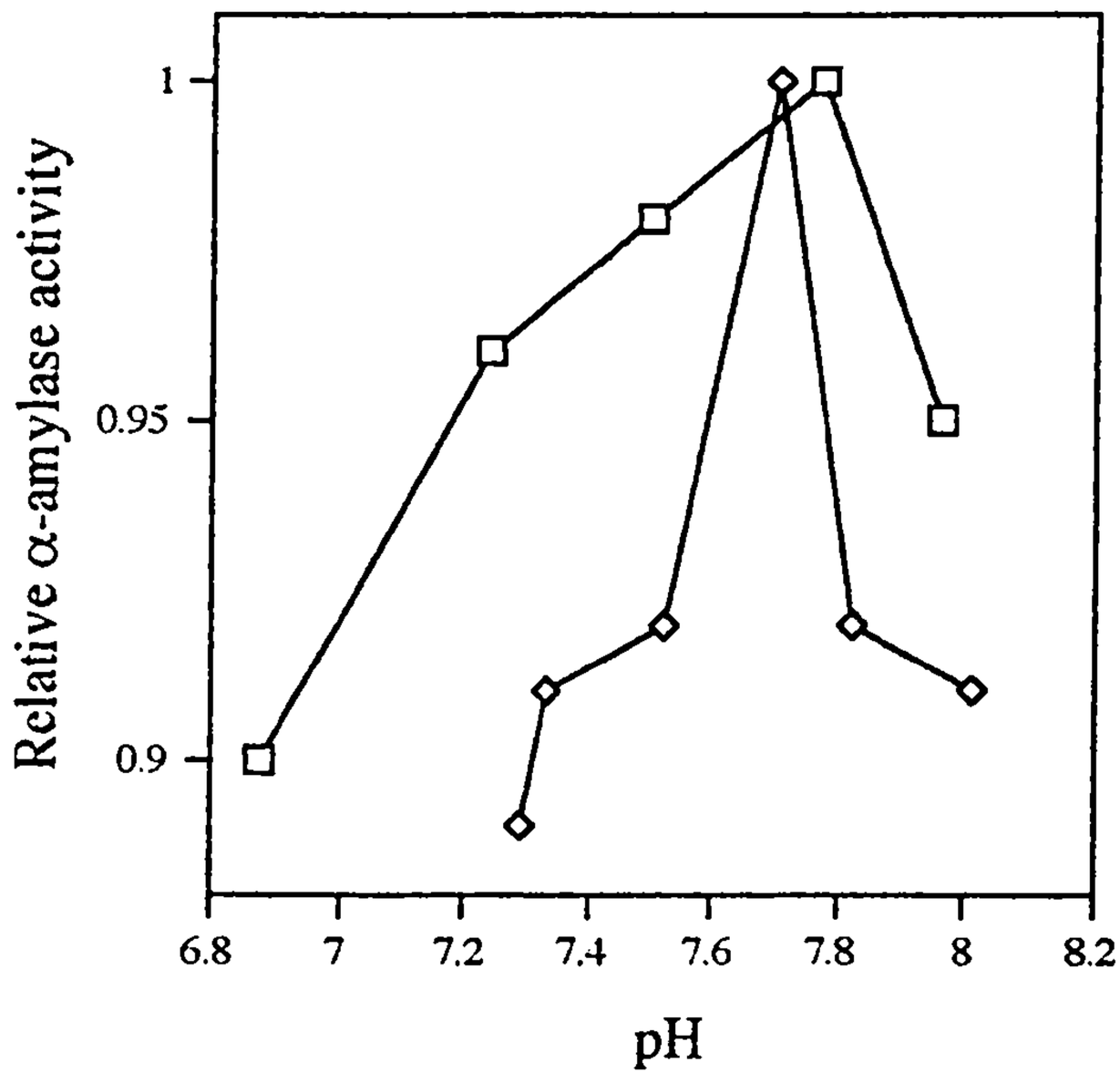


Fig. 7.3. pH optima of AmyLQS50.5 and AmyL.
Key:
—□— AmyL
—◇— AmyLQS50.5

7.2.2 Thermostability of AmyL and AmyLQS50.5

The thermostability of AmyL and AmyLQS50.5 was investigated by measuring the α-amylase activity remaining following incubation at 90°C (3.20). The activity of both enzymes decreased with time (Fig. 7.4), however, the rate of inactivation of AmyLQS50.5 was faster than that of AmyL. The reduced thermostability of AmyLQS50.5 when compared to AmyL was reflected in the half lives of irreversible thermoinactivation, which were 2 and 14 min, respectively.

The value obtained in this study for the half life of AmyL at 90°C was less than reported previously (Suzuki *et al.*, 1989) which was tested in the presence of 10 mM $CaCl_2$. Calcium is known to increase the thermostability of AmyL (Violet and Meunier, 1989; Machius *et al.*, 1995) and it is likely that the lower concentration of calcium used in this study was responsible for the observed reduction in half life. The affect of calcium concentration on the thermostability of AmyL and AmyLQS50.5 was not investigated.

Irreversible thermoinactivation of AmyL involves the deamidation of asparagine and/or glutamine residues (Tomazic and Klibanov, 1988). This differs from that of AmyQ and AmyS which proceeds *via* monomolecular conformational scrambling. The degree of deamidation of asparagine and glutamine residues is both sequence and conformation specific, with the former amino acid more susceptible to deamidation than the latter. The amino acid sequence of AmyLQS50.5 contains 23 asparagine and 15 glutamine residues. In contrast, AmyL has 25 asparagine and 20 glutamine residues. Reducing the number of amino acid residues susceptible to deamidation in AmyLQS50.5 did not cause a concomitant increase in the thermostability of this enzyme.

7.2.3 Summary

When equivalent proteins from thermophilic and mesophilic bacteria are compared, a common feature of the latter proteins is the presence of hydrophobic alanine residues in place of basic lysine residues (Argos *et al.*, 1979). Such amino acid substitutions stabilize the structure of thermophilic proteins due to electrostatic interactions and the consequential affects of packing arrangements within the protein molecule. To engineer an α-amylase with an overall positive charge, basic amino acids were introduced into AmyL to increase the pI. Hence, it is possible that the electrostatic effects of basic amino acids on the structure of AmyLQS50.5 played a significant role in the reduced thermostability of this enzyme. The temperature and pH optima of AmyLQS50.5 were largely unaffected by the amino acid changes, suggesting that these properties were

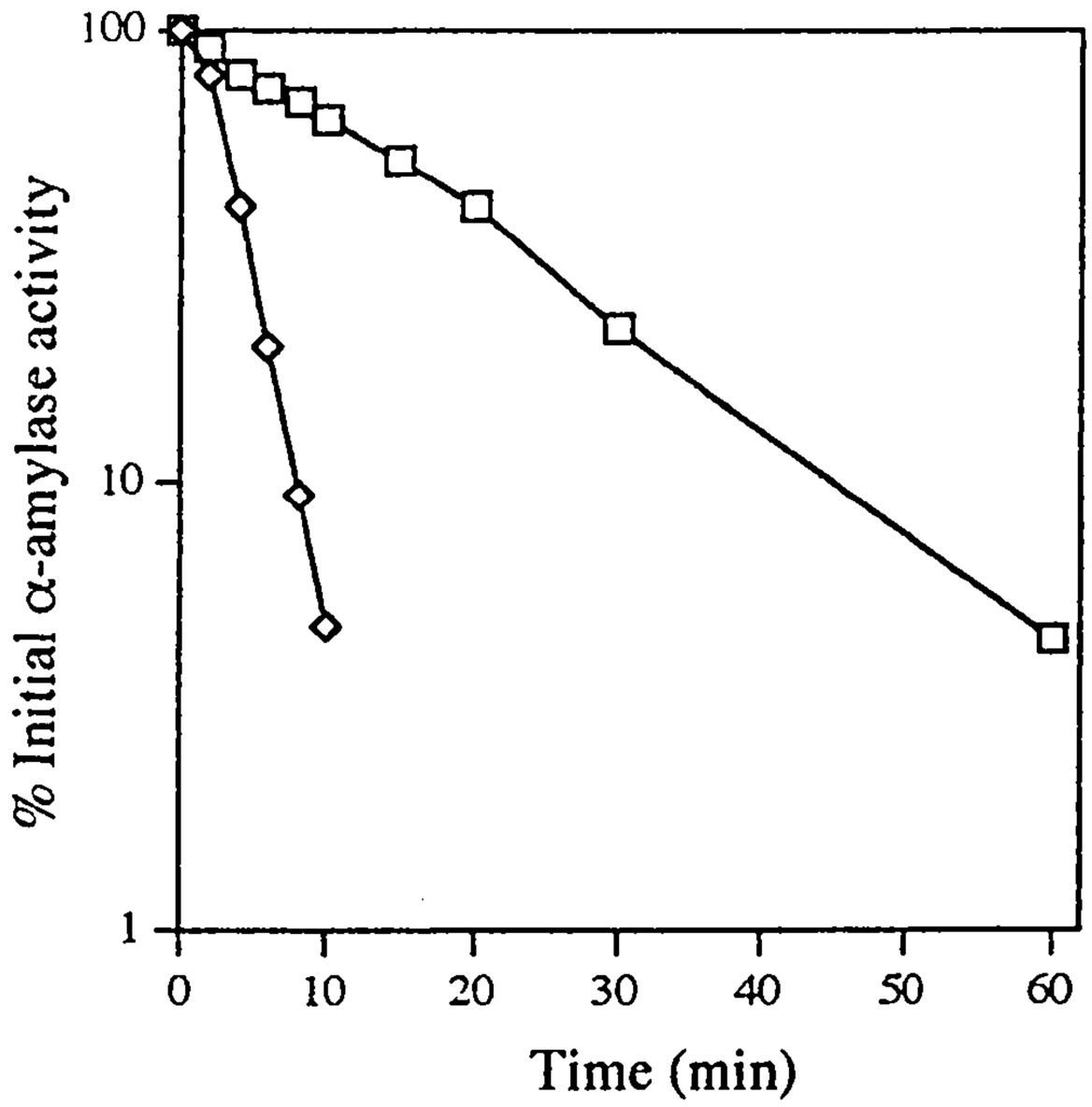


Fig. 7.4. Irreversible thermoinactivation of AmyLQS50.5 and AmyL in culture supernatants.
Key:
—□— AmyL
—◇— AmyLQS50.5

mediated by structural features of the protein other than those responsible for thermostability.

7.3 Factors affecting the stability of AmyLQS50.5 and AmyL

When compared with AmyL, the levels of AmyLQS50.5 detected in culture supernatants was reduced by approximately 4-fold (6.2). To determine whether this was the result of proteolytic degradation of AmyLQS50.5 in the supernatant, the stability of α-amylases in culture supernatants was investigated (3.21). The stability of the enzymes was investigated in the presence and absence of the Ca^{2+}chelator EDTA, to assess of influence of this ion in the maintenance of α-amylase structural integrity.

AmyL and AmyLQS50.5 were stable at 4°C for approximately 300 hours in the absence of EDTA (Fig. 7.5). The activity of AmyL appeared to fall to 75% of the initial level and then remained relatively constant. Similarly, the activity of AmyLQS50.5 fell initially to 85% of the starting value and then stabilized at this level. These observations were reflected by the immunoprecipitable bands present on the Western blots, the intensity of which did not decrease significantly with time.

In contrast, when EDTA was added to the supernatants, the activity of AmyL and AmyLQS50.5 decreased markedly. This was most pronounced with AmyLQS50.5; after approximately 300 hours at 4°C less than 10% of the initial activity remained. At the equivalent time point the activity of AmyL had decreased to 35%. Consistent with the activity data, the intensity of the α-amylase bands on the Western blots were also seen to decrease with time, confirming that the decrease in α-amylase activity in the presence of EDTA was caused by degradation and not enzyme inactivation.

B. subtilis elaborates several extracellular proteases and the culture supernatants used for these experiments would be expected to contain, in addition to α-amylases, enzymes with proteolytic activity. α-Amylase degradation products were observed on the Western blots in the presence of EDTA, the intensity of which increased with time (Fig. 7.5). These degradation products were most pronounced for AmyL, although equivalent AmyLQS50.5 degradation products were detectable. From the molecular size standards, the major degradation product of both α-amylases was estimated at 35 kDa.

After incubation for194 hours at 4°C, supernatant samples were removed and incubated at 37°C for 12 hours. Increasing the temperature from 4°C to 37°C in the presence of EDTA resulted in the complete proteolytic degradation of AmyLQS50.5 and AmyL, as indicated by Western blotting (Fig. 7.6) and a

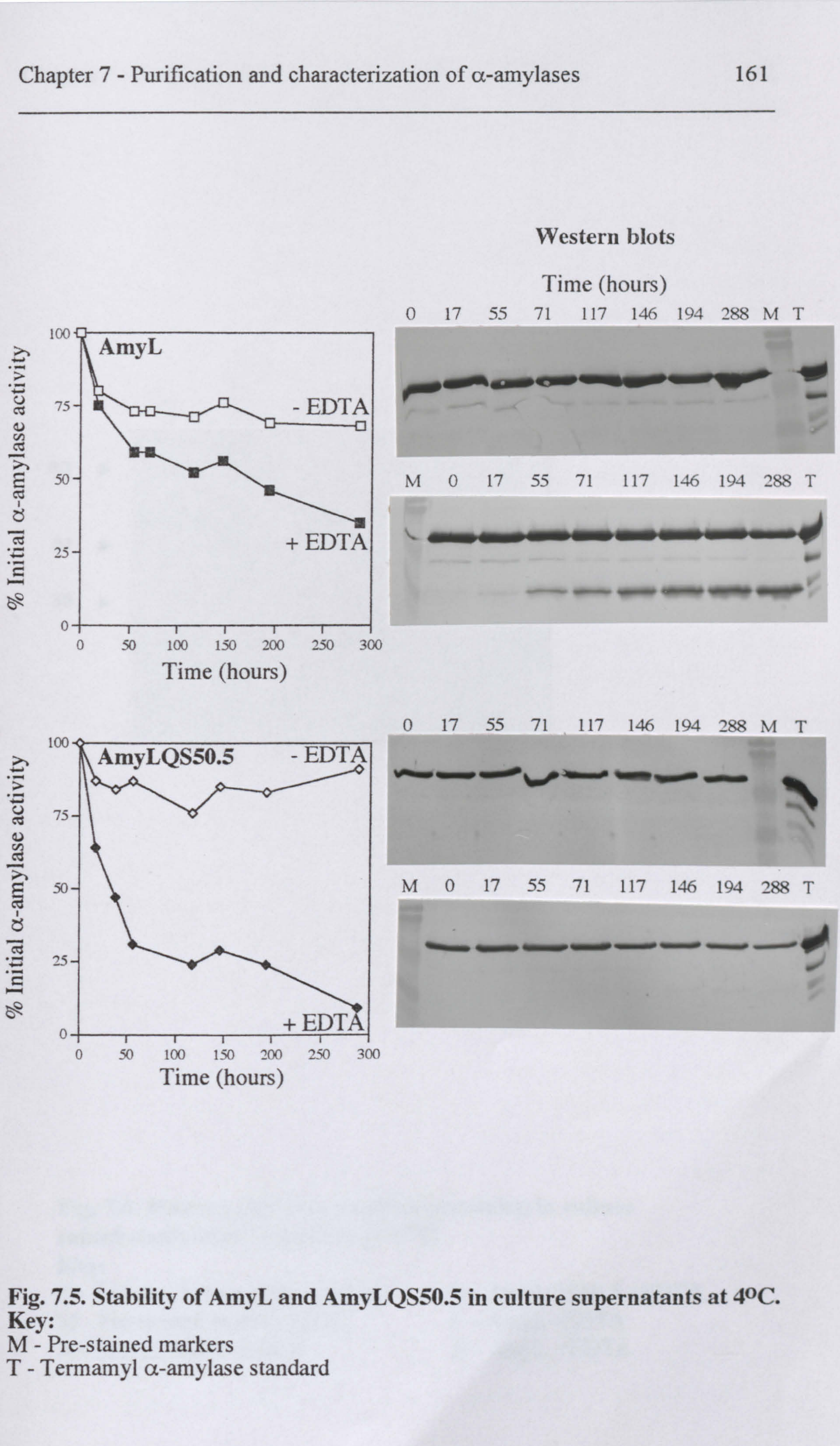


Fig. 7.5. Stability of AmyL and AmyLQS50.5 in culture supernatants at 4°C.
Key:
M - Pre-stained markers
T - Termamyl α-amylase standard

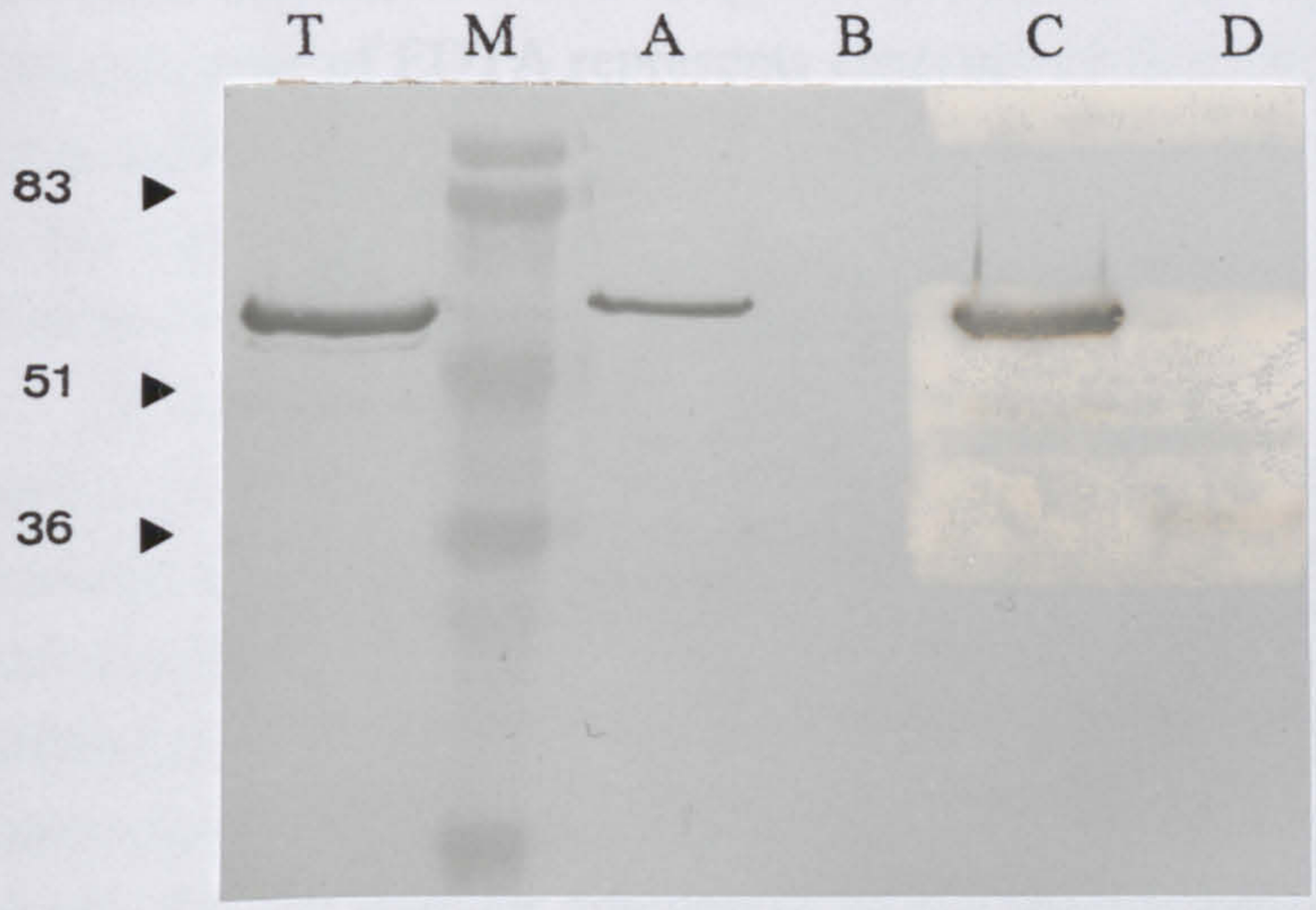


Fig. 7.6. Western blot of α-amylases remaining in culture supernatants after incubation at 37°C.

Key:

T - Termamyl α-amylase standard
M - Pre-stained markers (kDa)
A - AmyLQS50.5 -EDTA
B - AmyLQS50.5 +EDTA
C - AmyL -EDTA
D - AmyL +EDTA

lack of detectable α-amylase activity. In the absence of EDTA, the stability of both α-amylases was unaffected by the temperature shift.

It has been shown previously that, in the absence of Ca^{2+}, AmyL is susceptible to a protease produced by its native host, *B. licheniformis* (Machius *et al.*, 1995). The enzyme believed to be responsible, Glu-C endopeptidase, cleaves AmyL between amino acid residues 189 (glutamic acid) and 190 (asparagine) to generate products with calculated molecular masses of 21.7 kDa and 33.4 kDa. It is possible that the ~35 kDa AmyL/AmyLQS50.5 degradation products observed in the presence of EDTA represents cleavage of the α-amylases between residues 189 and 190. The production of an equivalent Glu-C endopeptidase by *B. subtilis* has not been reported and a search of the currently sequenced *B. subtilis* genes failed to identify a homologous protease.

In conclusion, the data presented here confirms the role of Ca^{2+} in the stability of α-amylases. In the presence of the chelator EDTA, AmyL and AmyLQS50.5 become susceptible to proteolysis, most likely as a consequence of structural disruptions in the absence of essential Ca^{2+} ions. At 4°C and 37°C, the stability of AmyLQS50.5 was affected most by the chelation of Ca^{2+}. This observation, together with the reduced thermostability of AmyLQS50.5 (7.2.2), indicate that the changes engineered to increase pI adversely affected the structure of this α-amylase. When Ca^{2+} ions were not chelated, both enzymes were stable and retained enzymatic activity.

7.4 Integration of α-amylase genes into the chromosome of *B. subtilis*

In the absence of selection, plasmids based on pUB110 are known to be structurally and segregationally unstable in *B. subtilis* (Gryczan, 1982; Bron and Luxen, 1985). To avoid problems associated with plasmid instability, pKS401 was constructed to facilitate the integration of xylose-inducible α-amylase genes into the chromosome of *B. subtilis* DN1885 (4.4). Plasmid pKS401 has an *amyL* gene than can be replaced with the constructed chimeric α-amylase genes. The *Pst*I to *Bam*HI *amyL* fragment from pKS401 was replaced with the corresponding fragments from pKS302, pKS304, pKS305 and pKS308 using *E. coli* XL1-Blue as the host. The resulting integration plasmids, with their respective α-amylase gene constructs, are shown in Table 5.19.

Purified integration plasmids were used to transform naturally competent *B. subtilis* DN1885 (3.9) selecting for transformants on agar plates containing chloramphenicol. Integration of pKS401-based plasmids into the chromosome of *B. subtilis* DN1885 was achieved by homologous recombination between plasmid- and chromosomally-located *xylR* genes, resulting in the duplication of this gene (Fig. 7.8). Integrants were checked for the ability to synthesize α-

amylase using the plate assay (Fig. 7.9). All strains except DN1885 *xylR*::pKS404 produced α-amylase in the presence of xylose. This was expected since AmyLQS50 could not be detected in the supernatant of DN1885 (pKS304) (6.2, 6.3).

None of the strains produced zones of starch hydrolysis in the absence of xylose, confirming the repression of α-amylase synthesis in the absence of inducing substrates. In the absence of xylose DN1885 (pKS305) was still capable of synthesizing large quantities of the wild type α-amylase (6.2). However, when cloned in pKS408 and integrated into the chromosome, the same α-amylase gene regained the inducible phenotype. This confirms that the high level of α-amylase production observed in DN1885 (pKS305) was caused by a mutation and/or rearrangement located outside of the coding region for the mature form of AmyL, since this was the only part of the gene which was encoded by pKS408.

7.4.1 Southern hybridization

To confirm that the plasmids had been integrated at the *xyl* locus of the *B. subtilis* DN1885 chromosome, Southern hybridization was carried out on *Bgl*II digested chromosomal DNA (3.22) from *B. subtilis* DN1885, DN1885 *xylR*::pKS402, DN1885 *xylR*::pKS404, DN1885 *xylR*::pKS405B and DN1885 *xylR*::pKS408 (3.23), with the *xylR* gene as the hybridization probe.

Within the *xyl* regulon there are three *Bgl*II restriction sites (Hastrup, 1988); one upstream of *xynC*, a second within the *xylA* gene and the third at the start of *xylB* (Fig. 7.8). There are no *Bgl*II sites within the integration plasmids and therefore, once integrated, the size of the *xylR* chromosomal *Bgl*II fragment increases by approximately 9 kbp.

When chromosomal DNA was probed with *xylR*, a single DNA fragment produced a positive hybridization signal in each case (Fig. 7.10). The hybridizing fragment in *B. subtilis* DN1885 was much smaller (5.7 kbp from sequence data) than that of the other strains. The size of hybridizing fragment in the integrant strains was, as expected, approximately 15 kbp, confirming the integration of the plasmids at *xylR*.

Continued

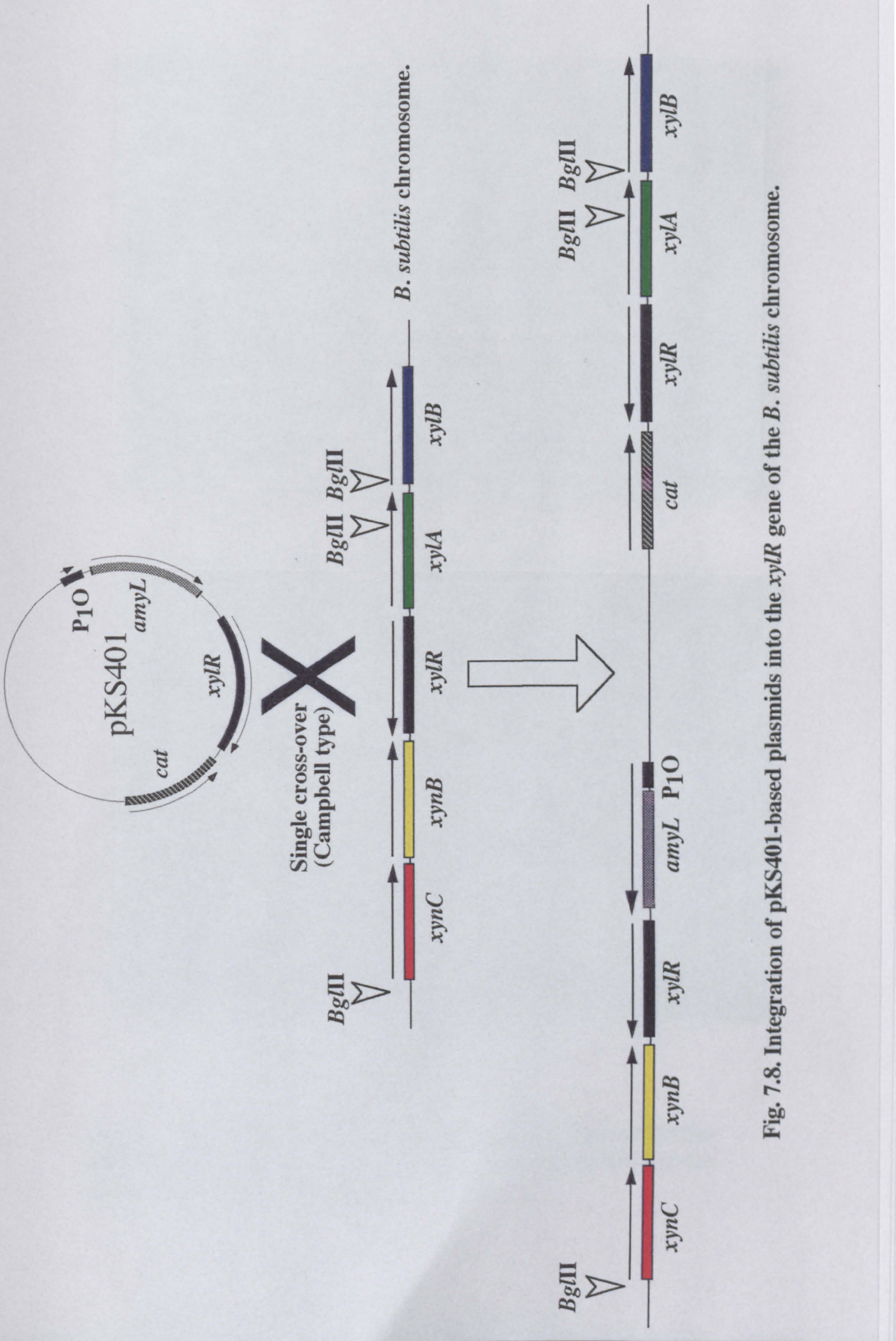
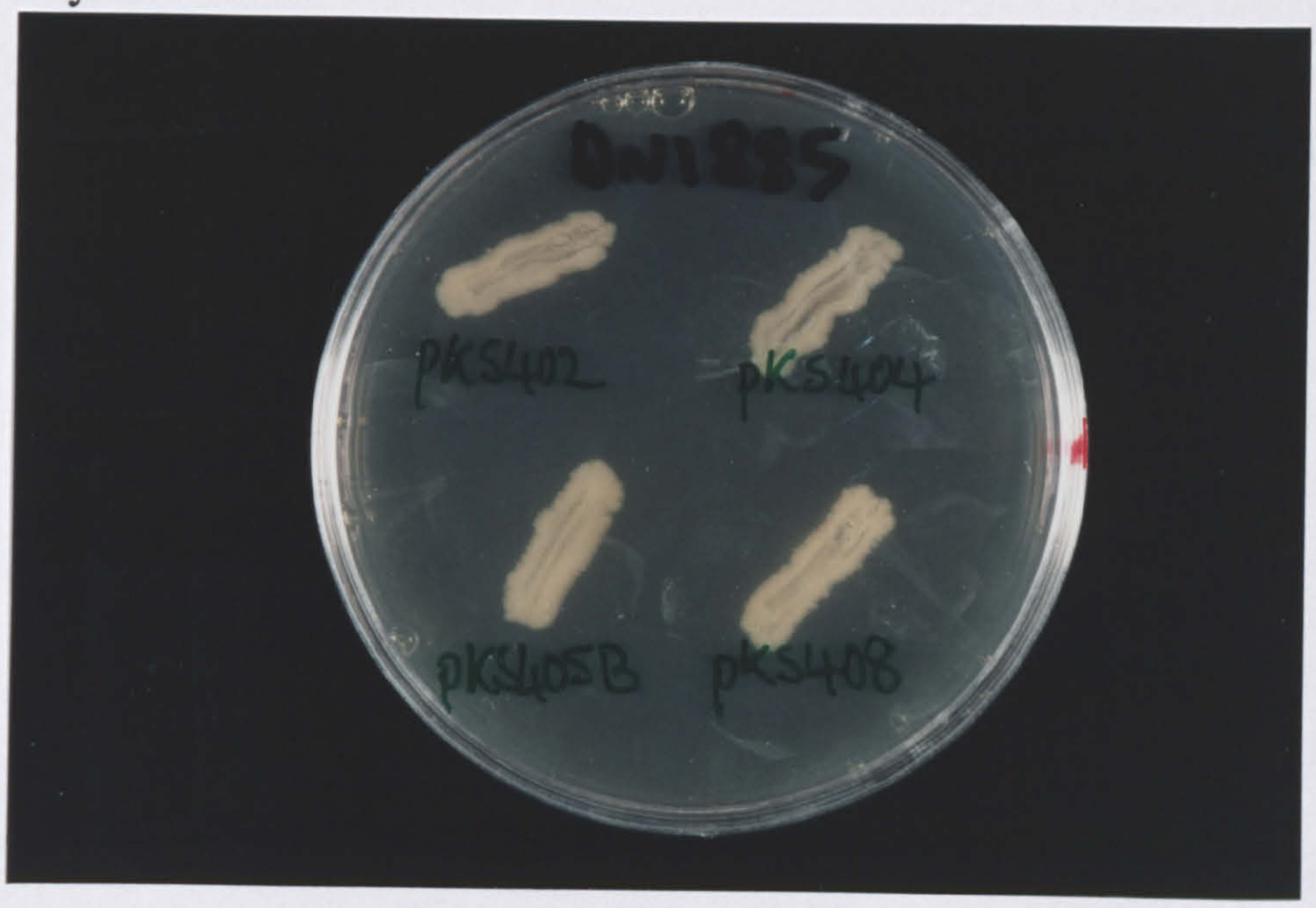


Fig. 7.8. Integration of pKS401-based plasmids into the *xylR* gene of the *B. subtilis* chromosome.

- Xylose



+ Xylose

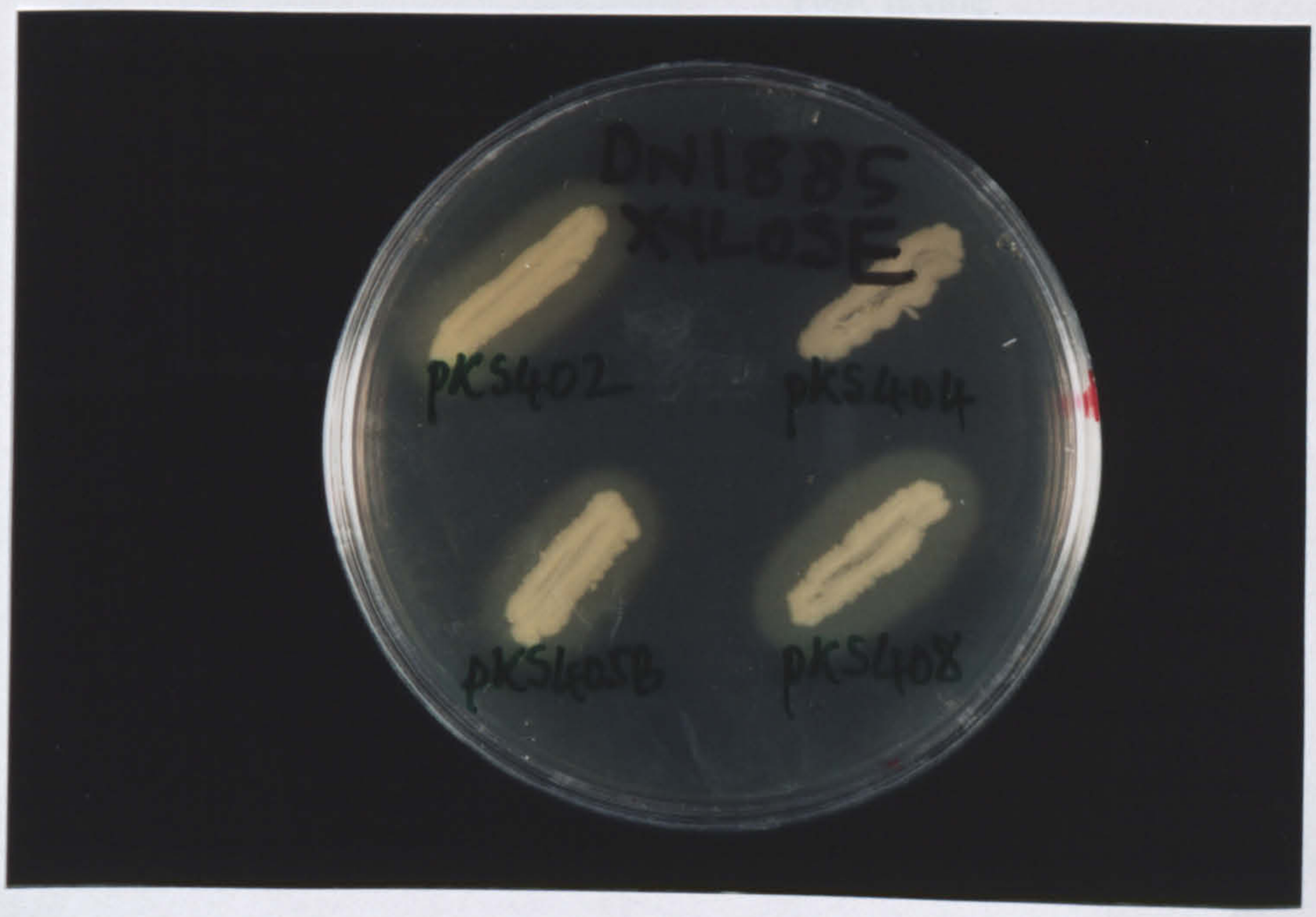
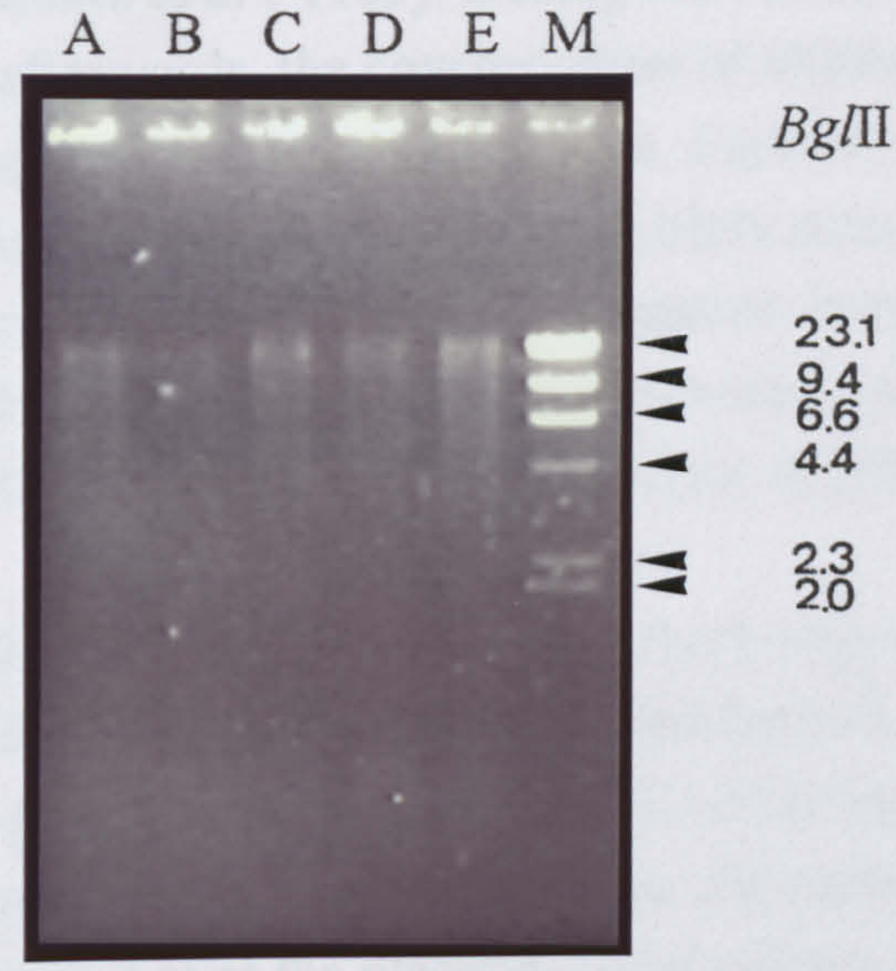


Fig. 7.9. Starch hydrolysis by derivatives of *B. subtilis* DN1885 with wild type and chimeric α-amylase genes integrated at the *xyl* locus of the chromosome.



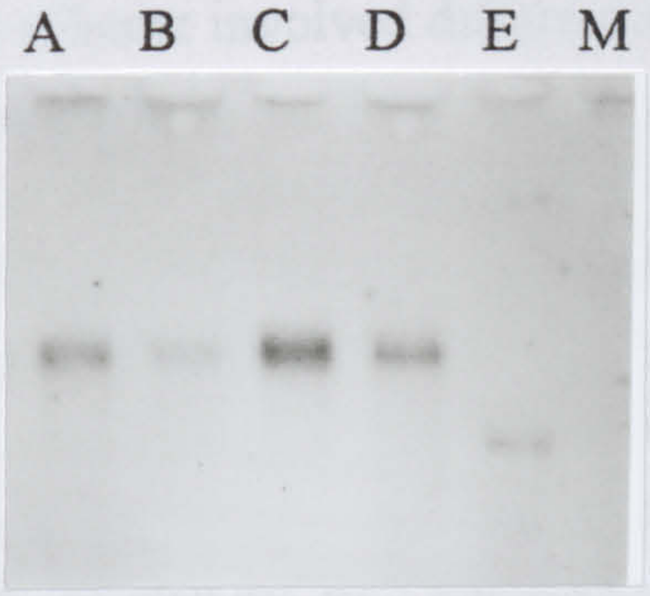


Fig. 7.10. Southern hybridization of derivatives of *B. subtilis* DN1885 with α-amylase genes integrated into the chromosome.
Key:
A - *B. subtilis* DN1885 *xylR*::pKS402
B - *B. subtilis* DN1885 *xylR*::pKS404
C - *B. subtilis* DN1885 *xylR*::pKS405B
D - *B. subtilis* DN1885 *xylR*::pKS408
E - *B. subtilis* DN1885
M - λ *Hind*III markers (kbp)

The copy number of plasmids integrated into the chromosome by single cross-over events can be increased by raising the concentration of antibiotic used for selection (Janniere *et al.*, 1985). During the selection for integration events, and at all stages afterwards, the concentration of chloramphenicol was maintained at 6 μg/ml to avoid amplification. Further, no evidence of amplification was visible from the Southern blots since hybridizing fragments of increased size would be expected in this situation. Integration of the α-amylase expression cassettes into the chromosome provided a stable system from which to study the secretion of α-amylases in the absence of selection.

7.5 Purification of α-amylases from culture supernatants

AmyLQS50.5 and AmyL were purified from *B. subtilis* DN1885 *xylR*::pKS405B and DN1885 *xylR*::pKS408 culture supernatants to allow a more detailed investigation of their properties. The chromosome-based expression systems were selected over the plasmid-based system for the large scale production of α-amylases to avoid the potential problem of plasmid instability. The purification scheme involved differential ammonium sulphate precipitation followed by separation on the basis of charge using the Rotofor cell (3.26).

Initially ammonium sulphate was added to culture supernatants to 50% saturation and precipitated protein was removed by centrifugation and resuspended in phosphate buffer. The concentration of ammonium sulphate in the supernatants was then increased to 80% saturation and newly precipitated protein removed by a second centrifugation step and resuspended in phosphate buffer. The proteins precipitated by 50% and 80% ammonium sulphate were analyzed by SDS-PAGE (3.15).

Addition of ammonium sulphate to 50% saturation resulted in the precipitation of AmyLQS50.5 and AmyL but relatively few other supernatant proteins (Fig. 7.11). In contrast, the addition of ammonium sulphate to 80% saturation precipitated, in addition to the α-amylases, a large number of other proteins. To facilitate subsequent purification steps only the protein samples precipitated by 50% ammonium sulphate were selected for separation using the Rotofor cell.

The 50% samples were separated using the Rotofor and the resultant fractions were analyzed by SDS-PAGE and assayed for α-amylase activity (3.14). When the precipitated proteins from DN1885 *xylR*::pKS408 were focused, AmyL was detected in fractions 5 to 19 (Fig. 7.12 A, B), with the peak of activity being in fraction 12 (Fig. 7.14). Apart from AmyL, no other protein bands were visible in this fraction; contaminating proteins were restricted to

fractions 1 to 6. Low molecular weight ampholytes were also visible on the gels in fractions 8 to 20.

When the precipitated proteins from DN1885 *xylR*::pKS405B were subjected to Rotofor separation, AmyLQS50.5 was detected in fractions 15 to 20 (Fig. 7.13 A, B) and fraction 19 was found to contain the most α-amylase activity (Fig. 7.14). As was the case for the protein sample from DN1885 *xylR*::pKS408, the majority of contaminating proteins and ampholytes were found in fractions 1 to 6 and 7 to 20, respectively.

In conclusion, AmyL and AmyLQS50.5 were separated from the majority of contaminating proteins present in the supernatants of their respective strains by the use of IEF. The peak α-amylase activities were found in fractions 12 (AmyL) and 19 (AmyLQS50.5) and these fractions were subjected to further purification steps. Following dialysis to remove ampholytes and concentration with polyethylene glycol, the purity of the α-amylase samples was confirmed by SDS-PAGE (Fig. 7.15).

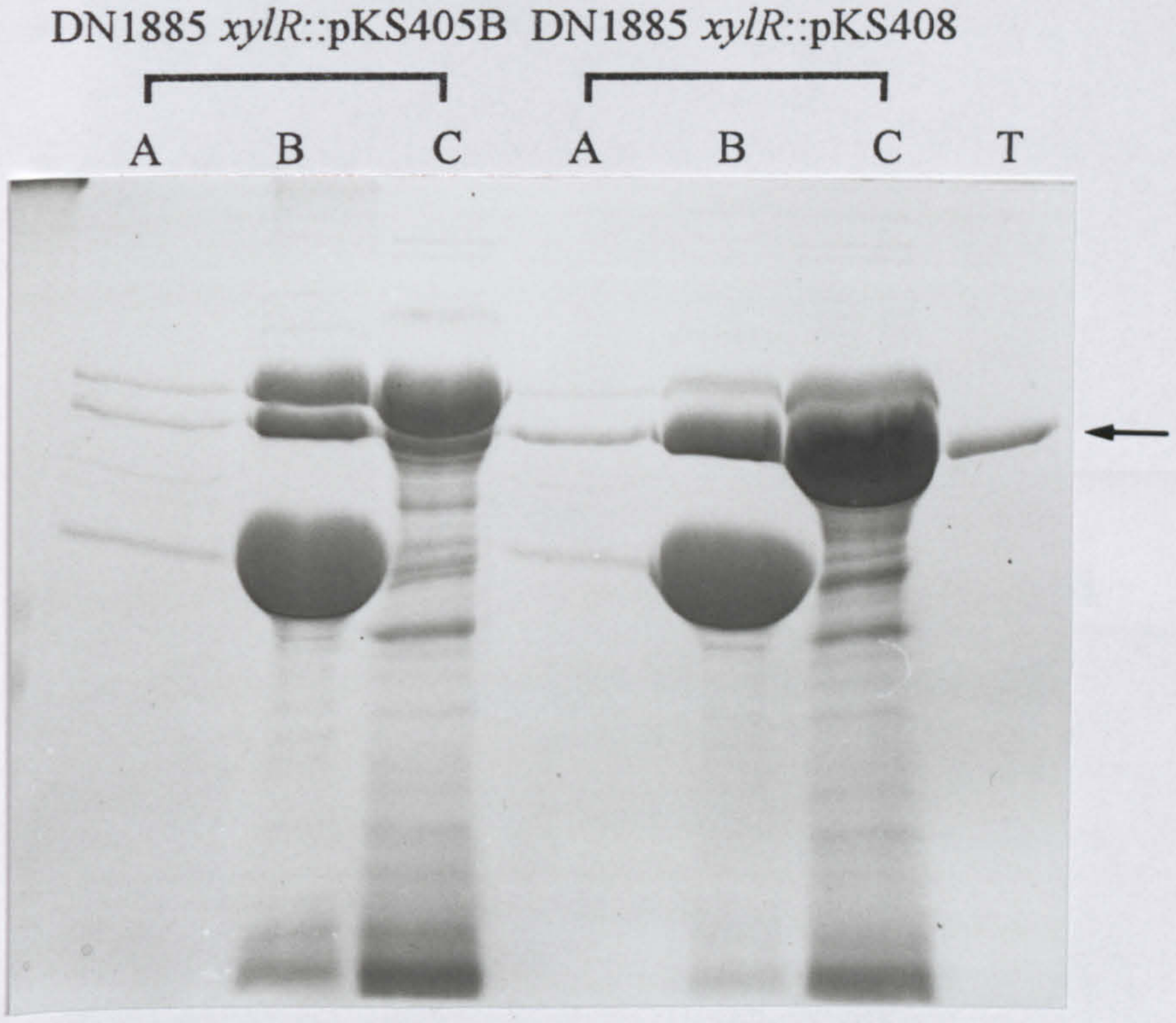


Fig. 7.11. SDS-PAGE of proteins precipitated from culture supernatants by the addition of ammonium sulphate.
Key:
A - Supernatant before the addition of ammonium sulphate
B - 50% saturation ammonium sulphate
C - 80% saturation ammonium sulphate
T - Termamyl α-amylase standard

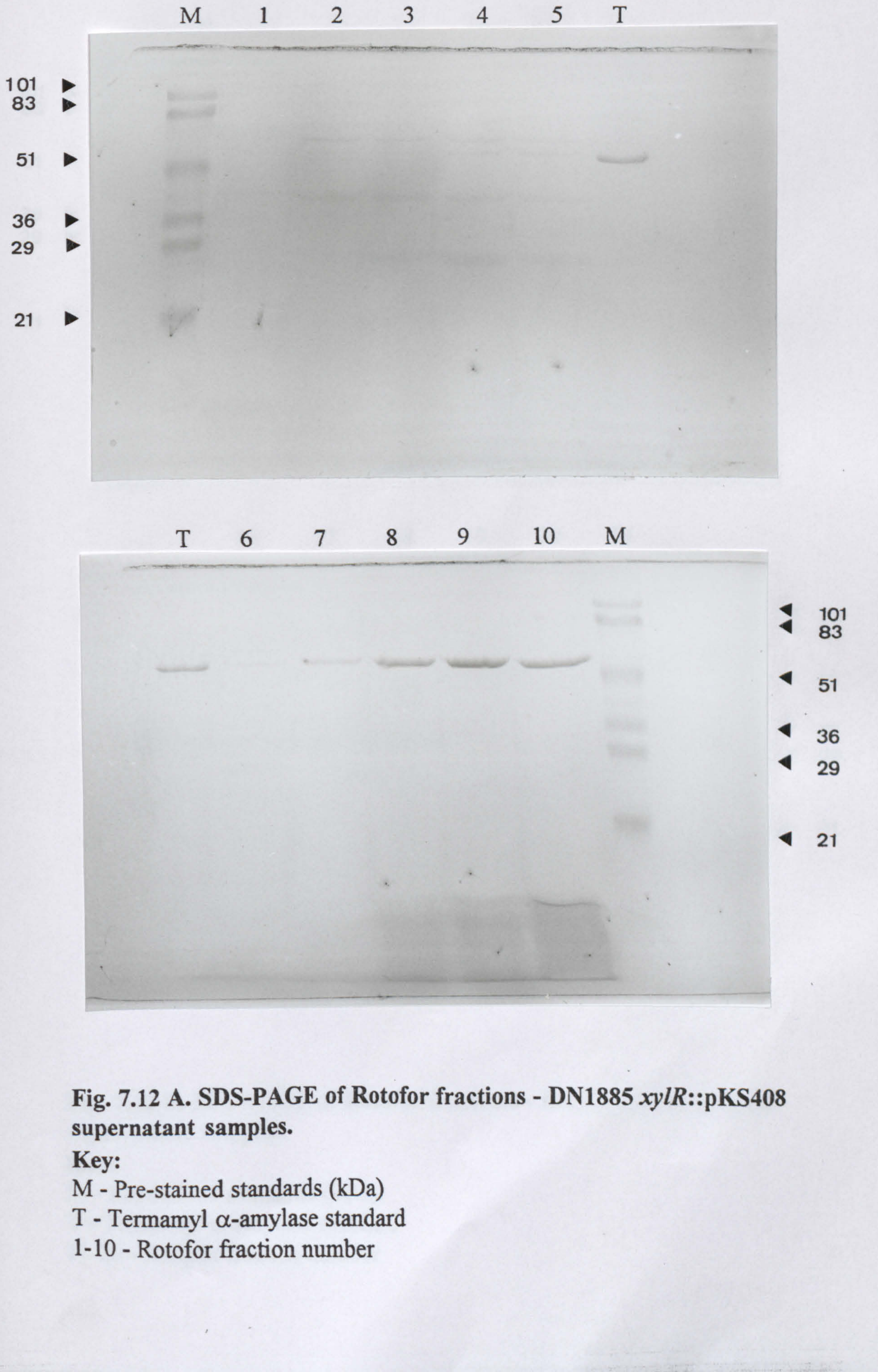


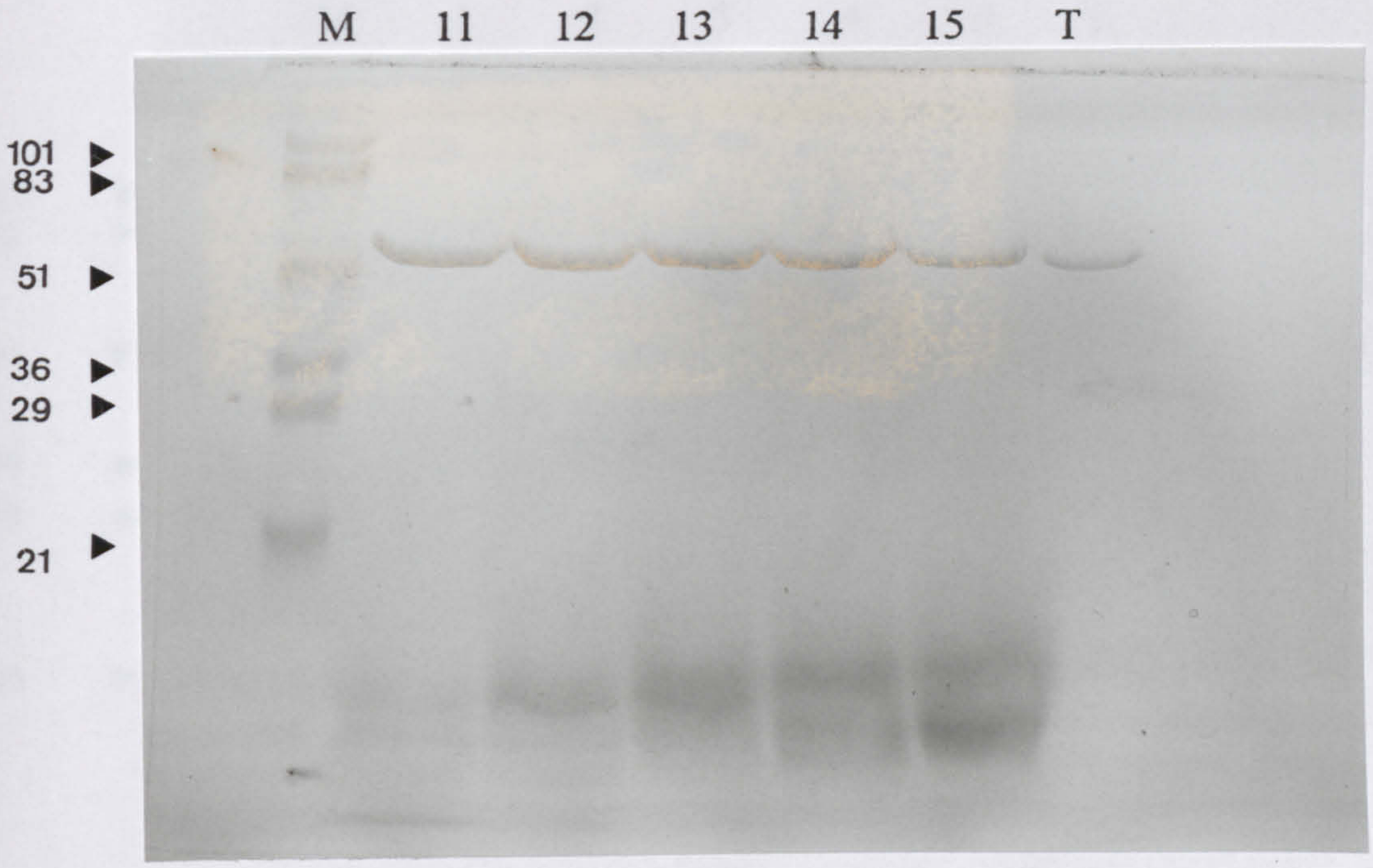
Fig. 7.12 A. SDS-PAGE of Rotofor fractions - DN1885 *xylR*::pKS408 supernatant samples.

Key:

M - Pre-stained standards (kDa)

T - Termamyl α-amylase standard

1-10 - Rotofor fraction number



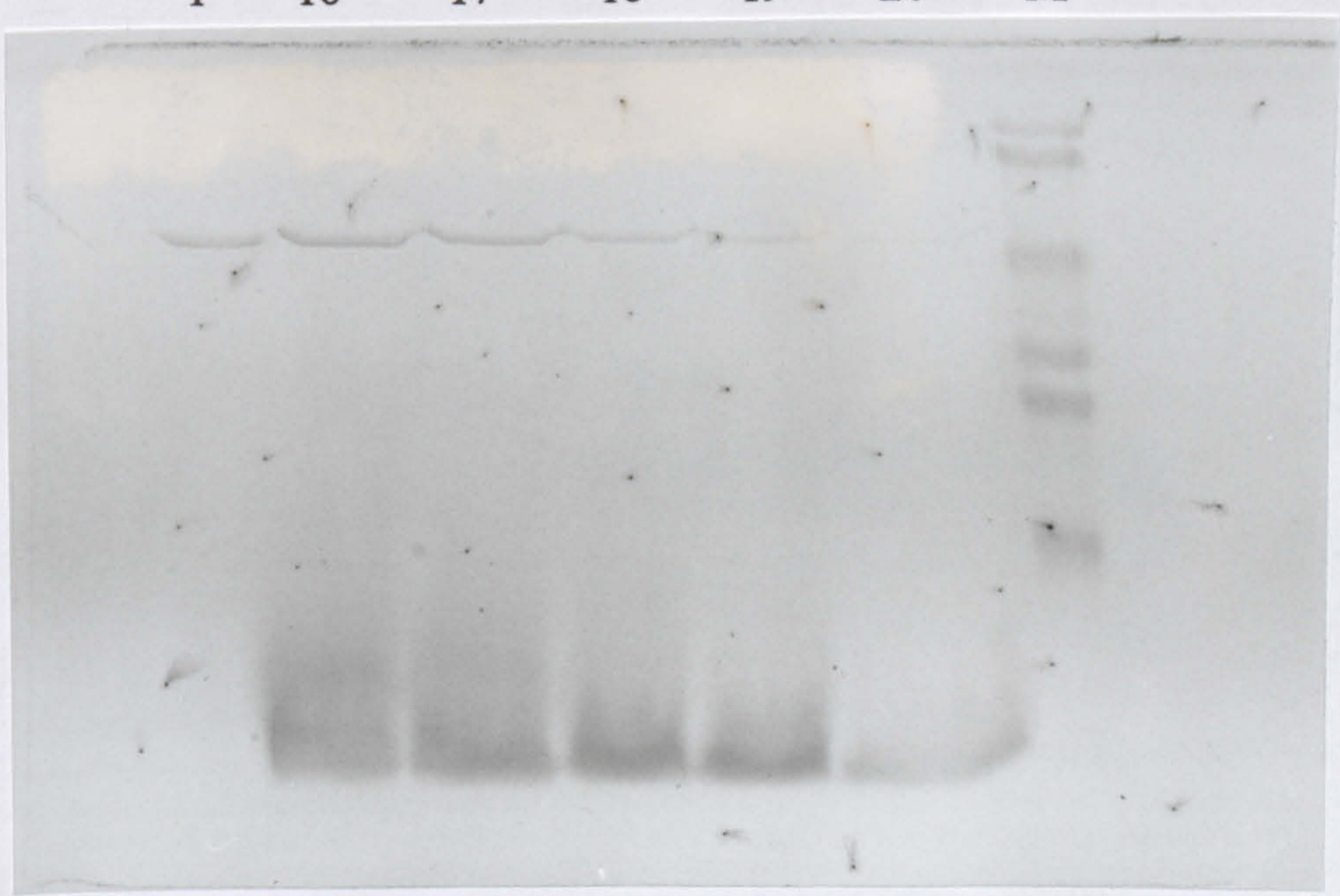


Fig. 7.12 B. SDS-PAGE of Rotofor fractions - DN1885 *xylR*::pKS408 supernatant samples.

Key:

M - Pre-stained standards (kDa)

T - Termamyl α-amylase standard

11-20 - Rotofor fraction number

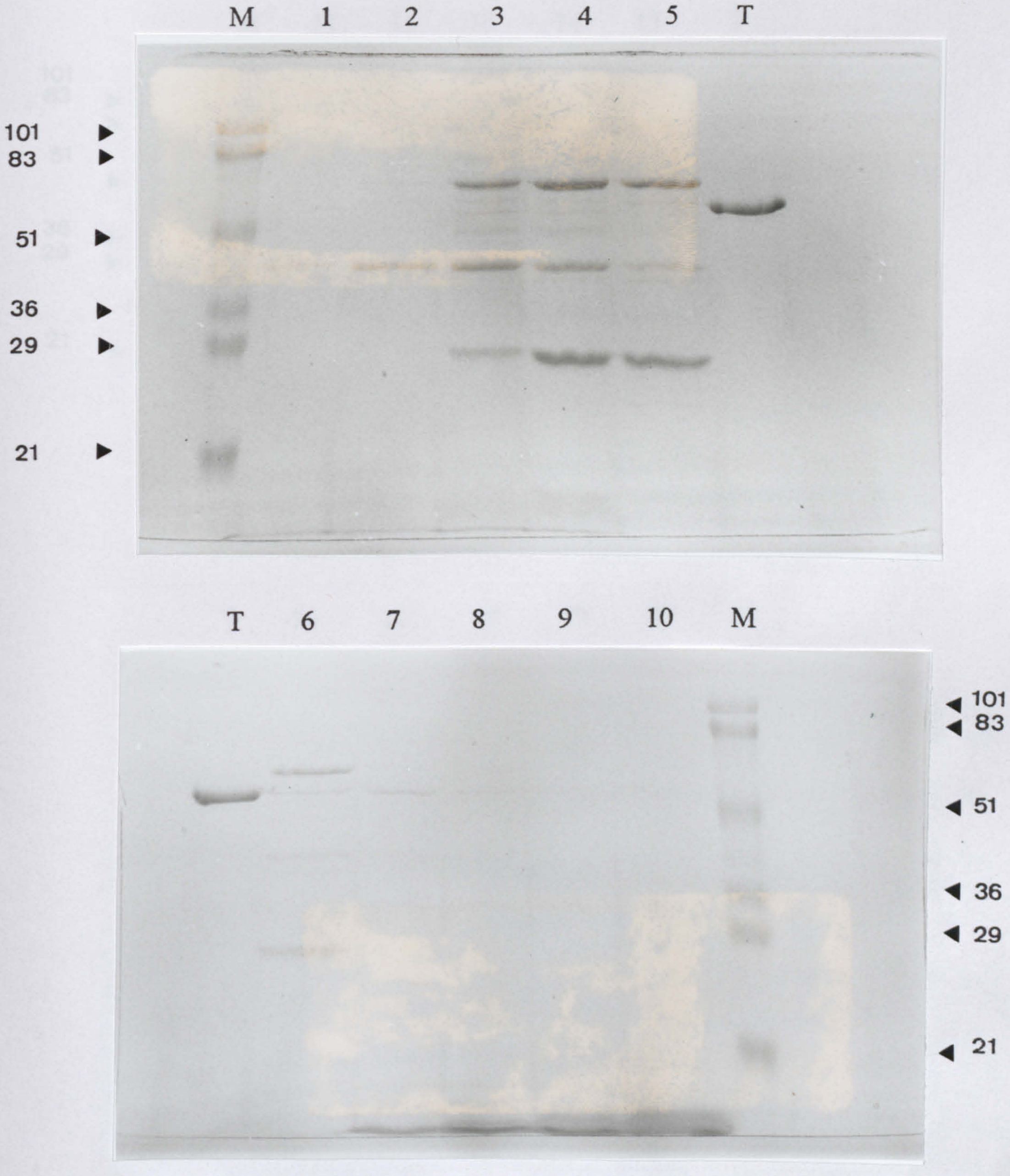


Fig. 7.12 A. SDS-PAGE of Rotofor fractions - DN1885 *xylR*::pKS405B supernatant samples.

Key:

M - Pre-stained standards (kDa)

T - Termamyl α-amylase standard

1-10 - Rotofor fraction number

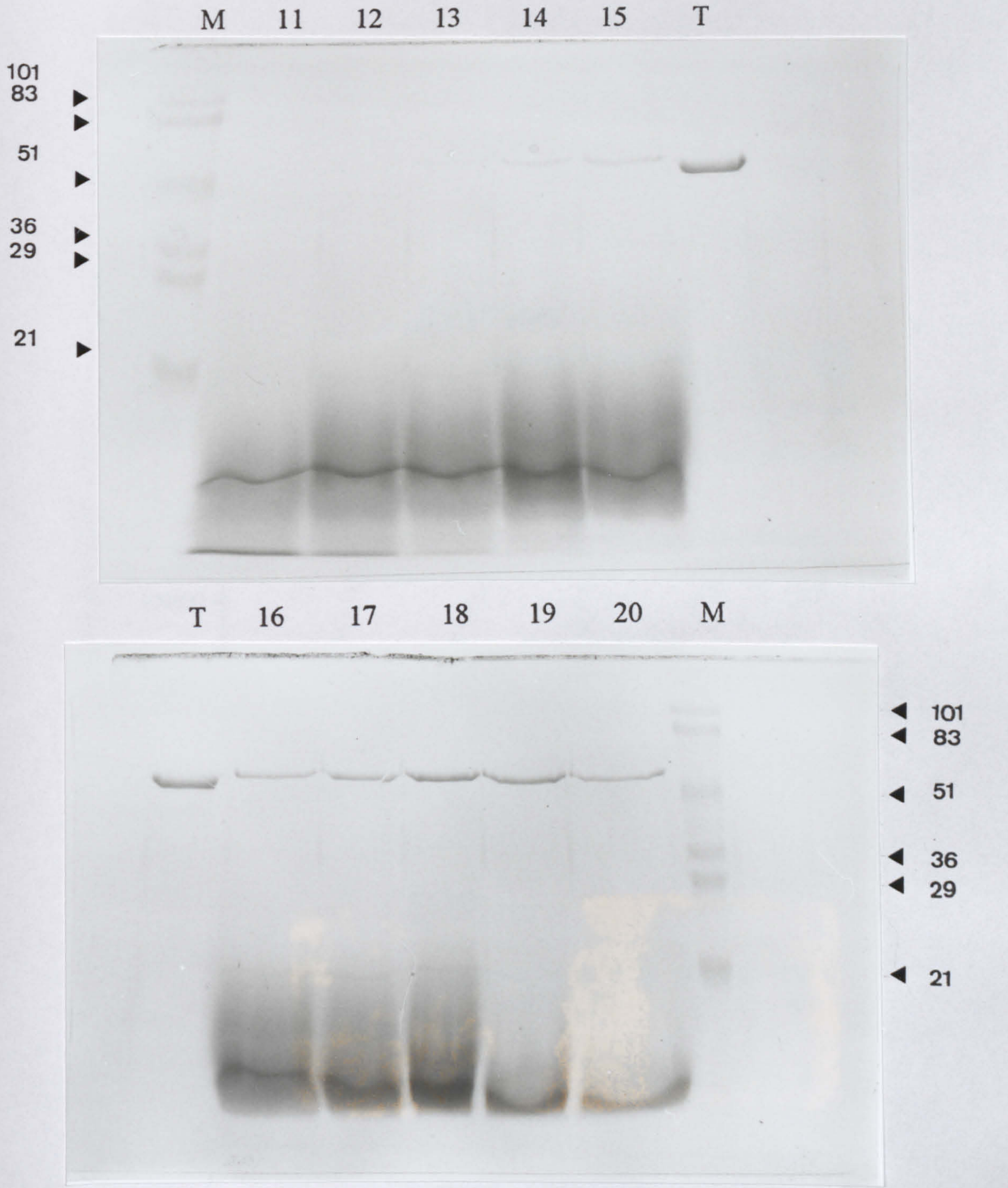


Fig. 7.12 B. SDS-PAGE of Rotofor fractions - DN1885 *xylR*::pKS405B supernatant samples.

Key:

M - Pre-stained standards (kDa)

T - Termamyl α-amylase standard

11-20 - Rotofor fraction number

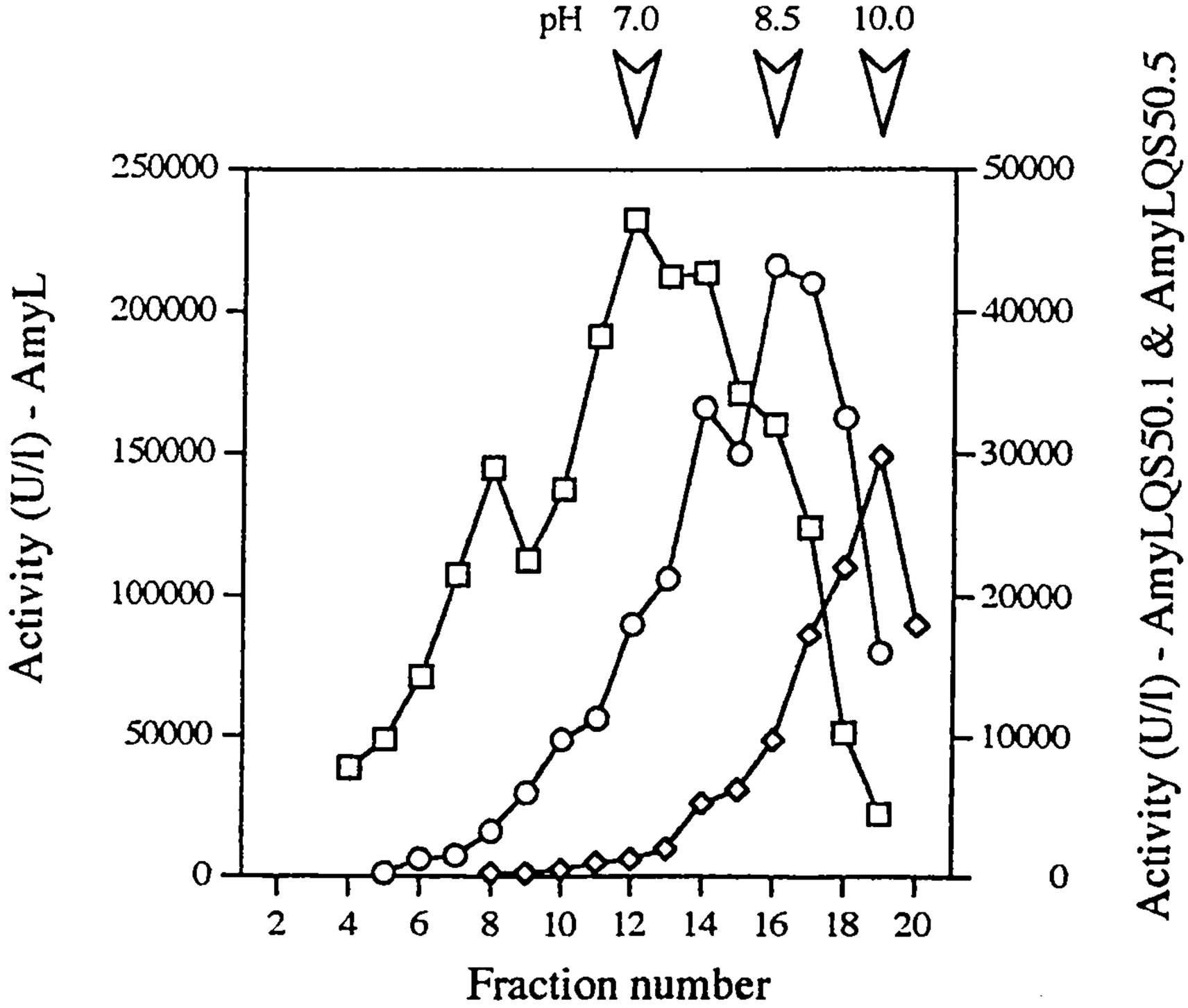


Fig. 7.14. Separation of precipitated supernatant proteins using the Rotofor - α-amylase activity of Rotofor fractions
Key:

—□— AmyL

—○— AmyLQS50.1

—◇— AmyLQS50.5

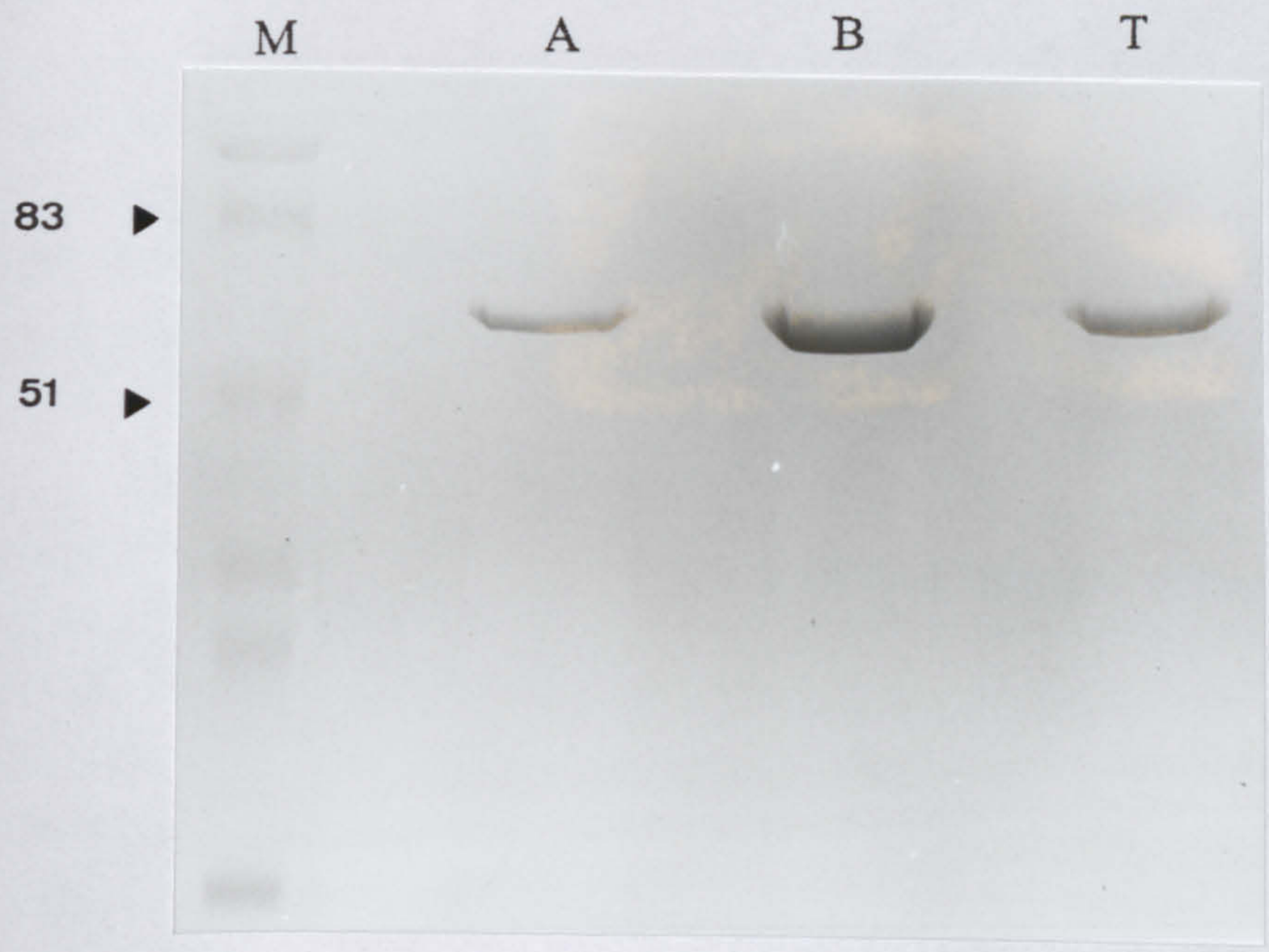


Fig. 7.15. SDS-PAGE of IEF-purified α-amylases.

Key:

M - Pre-stained standards (kDa)

A - AmyLQS50.5

B - AmyL

T - Termamyl α-amylase standard

7.6 Specific activities of AmyLQS50.5 and AmyL

To determine the affects of the engineered changes on AmyLQS50.5, the specific activity of this α-amylase was determined in relation to AmyL. Purified AmyLQS50.5 and AmyL (7.5) were diluted in 5 mM sodium phosphate buffer (pH 7.2) and their α-amylase activities (3.14) and protein concentration (3.27) determined. This data was used to calculate the specific activities of these enzymes (Table 7.2).

Table 7.2. The specific activities of purified AmyLQS50.5 and AmyL

α-Amylase	α-Amylase activity U/l	Protein concentration mg/l	Specific activity U/mg
AmyLQS50.5	9080.00	59.82	151.8
AmyL	38800.00	126.53	306.6

When compared to the wild type enzyme, there was a 2-fold reduction in the specific activity of AmyLQS50.5. Thus, although the temperature and pH optima of AmyLQS50.5 were largely unaffected by the engineered changes (7.2), the specific activity of the chimeric enzyme had been reduced by approximately 50%. This reduction in specific activity contributes in part to the lower levels of α-amylase activity detected in culture supernatants. Similar specific activity changes would be expected for other chimeric α-amylases, although this was not investigated.

7.7 Isoelectric points of AmyLQS50.5 and AmyL

The Rotofor was used to determine accurate pI values of α-amylases. AmyLQS50.1, AmyLQS50.5 and AmyL were precipitated from the supernatants of their respective strains by the addition of ammonium sulphate to 50% saturation (3.26). The protein samples were separated using the Rotofor cell and the α-amylase activities of the fractions determined (3.14).

The peaks of α-amylase activity for AmyL, AmyLQS50.1 and AmyLQS50.5 were found in fractions 12, 16 and 19, respectively (Fig. 7.14). The pH of these fractions was measured to within 0.5 pH units using pH test strips (Sigma). AmyL was found to have a pI of 7.0, a value which was consistent with that of rapidly exported *B. subtilis* proteins (Coxon, 1990). The pI of AmyLQS50.1 was determined as 8.5. This chimeric α-amylase was composed of AmyL amino acids 1 to 437 fused to residues 438 to 515 from of AmyS. Hence, the presence of 78 amino acids from the extreme C-terminus of the *B.*

stearothermophilus α-amylase increased the pI of the AmyLQS50.1 chimera by 1.5 units. Finally, the pI of the most extreme chimeric α-amylase, AmyLQS50.5, was measured at 10.0, a shift of 3 pH units when compared with AmyL.

In summary, the pI values of AmyL, AmyLQS50.1 and AmyLQS50.5 were determined by isoelectric focusing. In each case the actual pI values were higher than the value calculated using protein analysis software (Table 7.1). The pI of AmyL was found to be 7.0 and at physiological pH this α-amylase would possess little or no charge. This may represent a common evolutionary adaptation by members of the genus *Bacillus*, resulting in the production of secretory proteins with minimal potential for electrostatic interactions with the immobilized negative charge of the cell wall. The pI's of the chimeric α-amylases, AmyLQS50.1 and AmyLQS50.5, were increased with respect to AmyL and confirmed previous IEF data (7.1). On the basis of electrostatic interactions these positively charged α-amylases would be expected to interact with negatively charged groups in the cell wall-anionic polymer complex. The binding of AmyLQS50.5 and AmyL to the cell wall of *B. subtilis* was investigated and is discussed in section 8.3.

CHAPTER 8

FACTORS AFFECTING THE LATER STAGES OF THE SECRETION OF α-AMYLASES

8.1 The secretion of AmyL and AmyLQS50.5

When coupled with immunoprecipitation techniques, pulse-chase experiments can be used to study the processing and release of specific proteins. Pulse-chase and immunoprecipitation (3.28) was used to investigate the secretion of AmyL and AmyLQS50.5 from exponentially growing cultures of *B. subtilis* strains DN1885 *xylR*::pKS408 and DN1885 *xylR*::pKS405B, respectively. The latter strain was selected because, in the presence of xylose, it produced an active chimeric α-amylase with a pI significantly higher than AmyL. AmyLQS50.5 was also shown to cross-react with antiserum raised against AmyL (6.3).

Following pulse-labelling with ^{35}S methionine and immunoprecipitation, the precursor and mature forms of both α-amylases were visible on the autoradiographs and their relative mobilities were in agreement with their calculated molecular masses (Table 7.1). At time 0, approximately 80% of AmyL was in the mature form and by 1 min post-chase no precursor could be detected (Figs. 8.1 and 8.2). The processing of the AmyL precursor to the mature form was rapid and proceeded with a half life of ca. 15 sec. The amount of mature AmyL decreased steadily until 3 min post-chase, after which it remained constant at approximately 27% of the total AmyL (precursor + mature) present initially. Consistent with these observations, AmyL was released into the culture medium rapidly indicating that AmyL passed through the cell wall with little or no impedance.

Similarly, at time 0 approximately 72% of AmyLQS50.5 was in the mature form and by 4 min post-chase no precursor could be detected (Figs. 8.1 and 8.2). The rate of processing of the AmyLQS50.5 precursor was reduced when compared with AmyL, as reflected in a precursor half life of ca. 48 sec. The amount of mature AmyLQS50.5 increased at 30 sec to approximately 86% of the total, and then decreased steadily and dramatically with time until, at 5 min post-chase, no AmyLQS50.5 could be detected. Released AmyLQS50.5 could not be detected in the culture medium.

During or shortly after translocation, the mature forms of AmyL and AmyLQS50.5 were subjected to degradation, although the latter was affected to a much greater extent than the former. However, by the time the α-amylases reached the culture medium (only small amounts in the case of AmyLQS50.5), they were stable and retained catalytic activity over extended periods of time (7.3). This suggests that there is a time window between translocation and release into the culture medium during which the α-amylases are susceptible to proteolytic degradation.

Exoproteins fold into their native conformations on the outside of the cytoplasmic membrane either during the late stages of translocation or

immediately after exiting the translocation complex. Whilst in an unfolded or partially folded state, it appears that α-amylases are sensitive to the action of proteases. It was possible that the degradation of AmyL and AmyLQS50.5 observed during or following translocation was the result of the activity of proteolytic enzymes on incompletely folded α-amylases. It follows that if AmyLQS50.5 folds more slowly than AmyL this might account for the increased protease sensitivity of the chimeric enzyme.

The pulse-chase data, combined with data on the stability of the α-amylases (7.3) lead to the development of a working hypothesis that post-translocational folding of α-amylases was an important factor in the yield of secreted protein. The hypothesis proposes that exoproteins with reduced rates of folding are more susceptible to proteolytic degradation by either specific and/or non-specific proteases. We have started to test this hypothesis by determining the rate of folding of AmyL and AmyLQS50.5 *in vitro* (8.2).

In conclusion, pulse-chase experiments showed that both wild type and chimeric α-amylases were degraded during or shortly after translocation through the cytoplasmic membrane. AmyLQS50.5 was extensively degraded and consequently, it was not possible to determine whether the failure to detect this α-amylase in the culture medium was due to the extent of degradation or to binding to the cell wall (or a combination of the two). This degradation could account for the reduced levels of chimeric α-amylases in culture supernatants (6.2 and 6.3). Consistent with the data presented here, chimeric α-amylases with overall negative charge are also degraded in a similar manner (C L Jensen, personal communication).

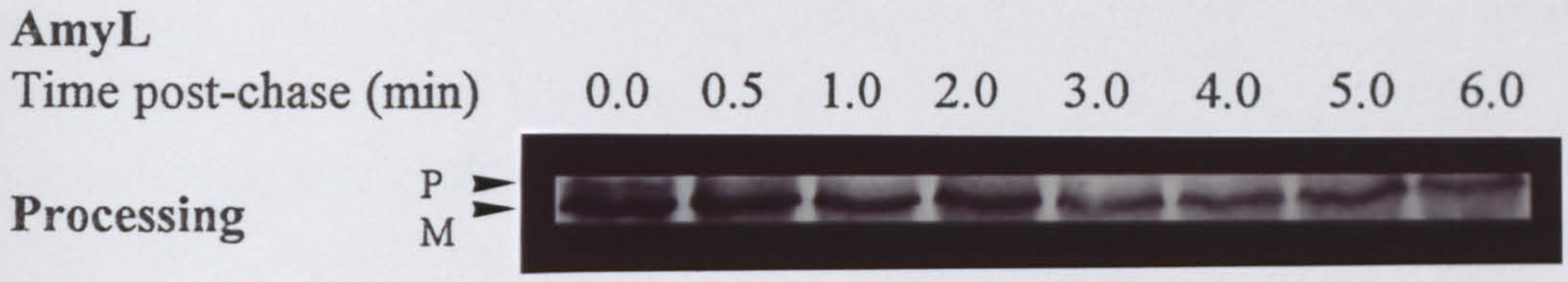


Fig. 8.1. Processing AmyL and AmyLQS50.5 and release into the supernatant.
Key:
P - Precursor protein
M - Mature protein

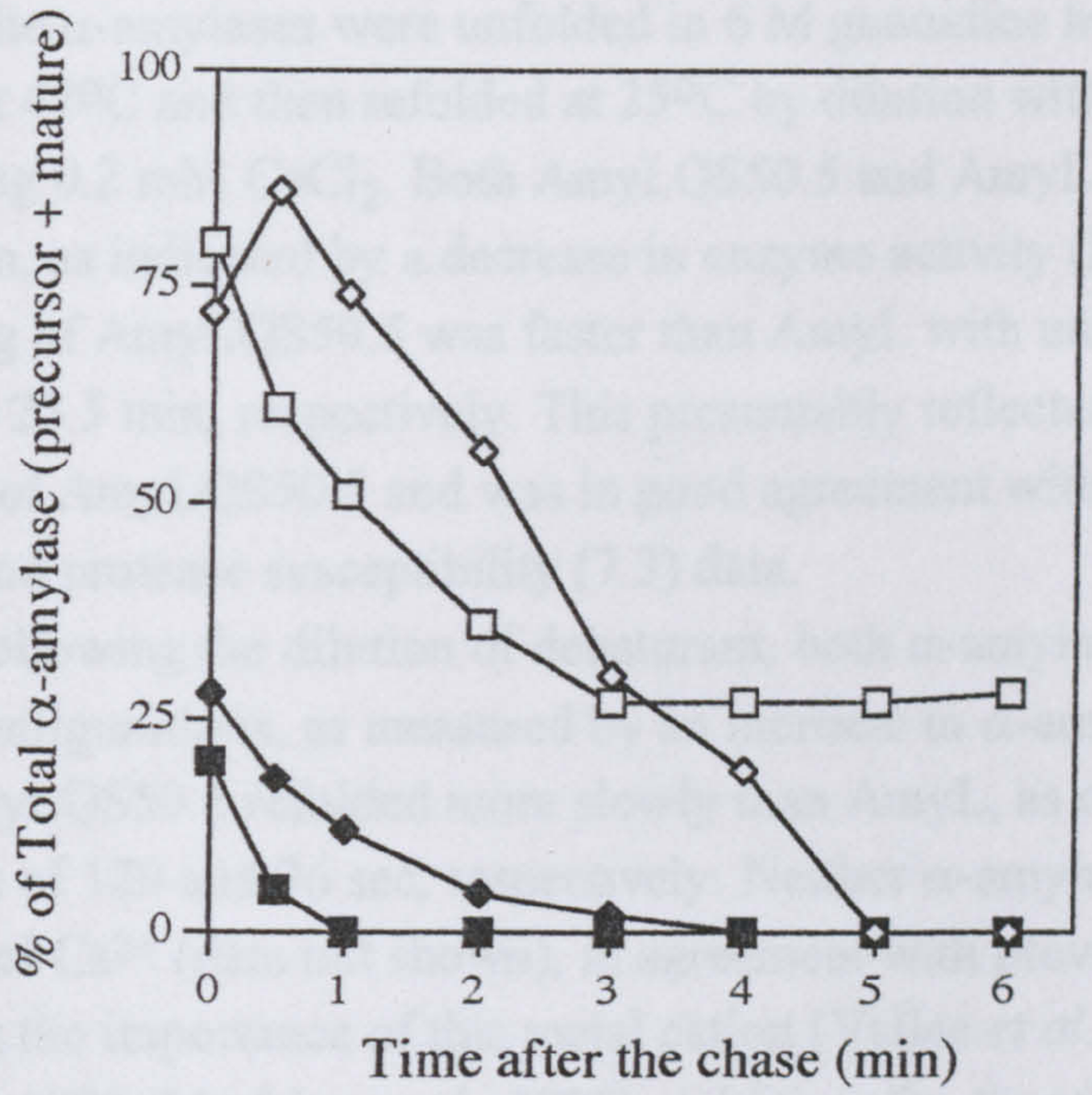


Fig. 8.2. Relative amounts of precursor and mature forms of AmyL and AmyLQS50.5.
Key:

—□— Mature AmyL

—◇— Mature AmyLQS50.5

—■— AmyL precursor

—◆— AmyLQS50.5 precursor

8.2 Unfolding/refolding transition of AmyLQS50.5 and AmyL

Several lines of evidence point to the rate of co-/post-translocational folding being an important factor in the increased susceptibility of AmyLQS50.5 to proteases during export. To test this hypothesis, the unfolding/refolding transitions of purified AmyLQS50.5 and AmyL (7.5) were investigated using the resistance of native α-amylases to proteolysis as a measure of the conformational state (3.29).

The α-amylases were unfolded in 6 M guanidine hydrochloride/1 mM EDTA at 47°C and then refolded at 25°C by dilution with acetate buffer containing 0.2 mM $CaCl_2$. Both AmyLQS50.5 and AmyL unfolded over a period of 90 min, as indicated by a decrease in enzyme activity (Fig. 8.3). The rate of unfolding of AmyLQS50.5 was faster than AmyL with unfolding half lives of 10.3 and 25.5 min, respectively. This presumably reflected reduced structural stability of AmyLQS50.5 and was in good agreement with the thermostability (7.2.2) and protease susceptibility (7.3) data.

Following the dilution of denaturant, both α-amylases refolded into their native configurations, as measured by an increase in α-amylase activity (Fig. 8.3). AmyLQS50.5 refolded more slowly than AmyL, as reflected in refolding half lives of 120 and 36 sec, respectively. Neither α-amylase refolded in the absence of Ca^{2+} (data not shown), in agreement with previous structural studies revealing the importance of this metal cation (Vallee *et al.*, 1959; Violet and Meunier, 1989; Machius *et al.*, 1995). Additionally, the yield of refolding of the two enzymes, as determined from the unfolding/refolding data, was 57% for AmyLQS50.5 and 81% for AmyL.

In conclusion, the data from *in vitro* protein folding experiments was in agreement with our working hypothesis that the rate of folding at the outer surface of the cytoplasmic membrane is an important component of the secretion pathway. The data showed that the rate of refolding of AmyLQS50.5 was reduced by approximately 3-fold when compared with AmyL.

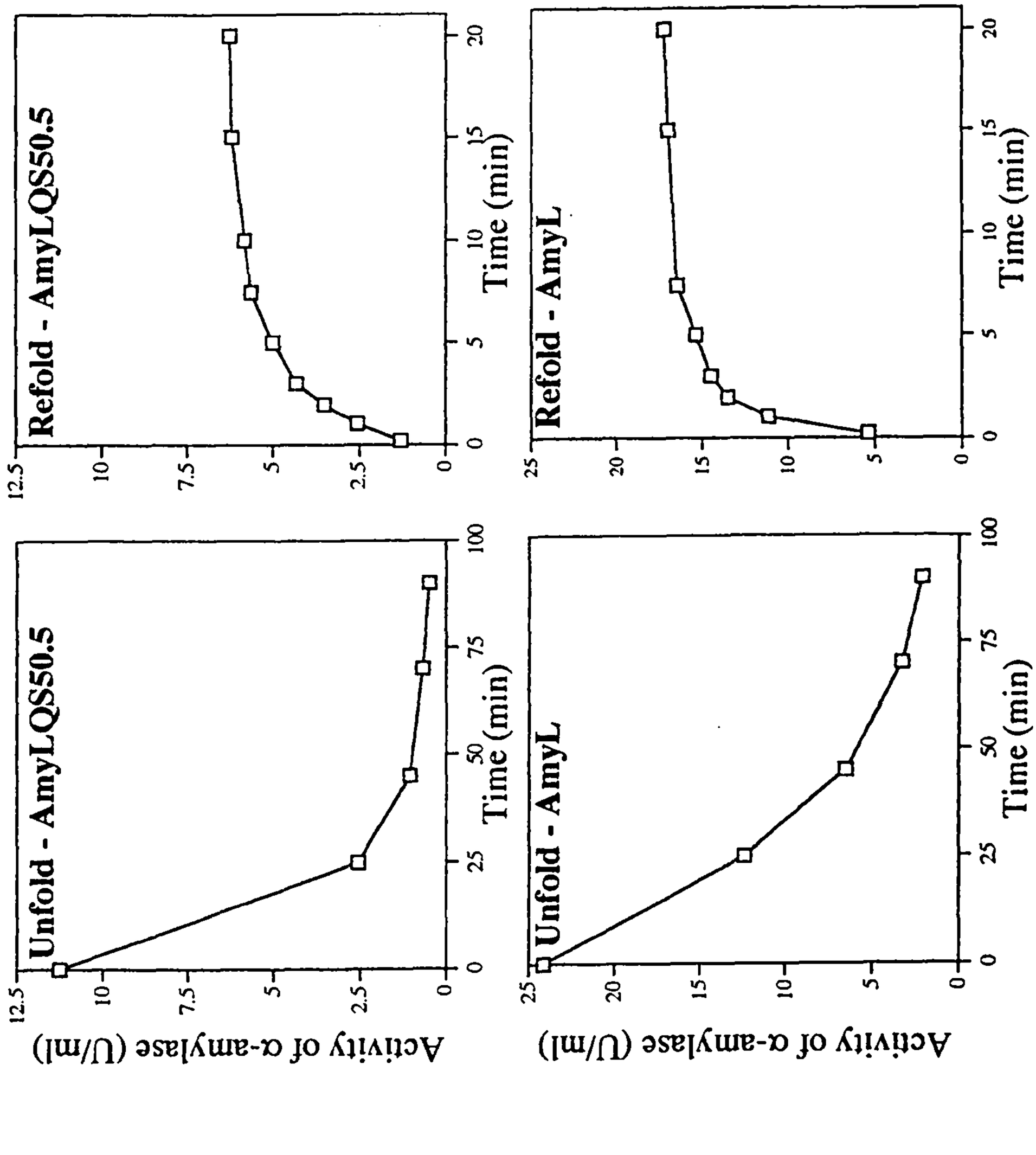


Fig. 8.3. The unfolding/refolding transitions of AmyLQS50.5 and AmyL.

8.3 Interaction of proteins with the *B. subtilis* cell wall

On the basis of its net positive charge, AmyLQS50.5 would be expected to bind to anionic groups within the *B. subtilis* cell wall after translocation through the cytoplasmic membrane. Since this effect could not be readily demonstrated *in vivo* (8.1), an *in vitro* cell wall binding (IVWB) assay was developed (3.30).

The IVWB assay was used to assess the ability of AmyLQS50.5 and AmyL to bind to isolated cell walls whilst in their native (fully folded) conformations. Further, to evaluate the potential effect of the cell wall on the secretion of a heterologous protein, the binding of HSA to walls was also investigated. HSA was chosen because it has been suggested previously that the secretion of this eukaryotic protein from *B. subtilis* is blocked by the cell wall (Saunders *et al.*, 1987).

Following interaction between purified proteins and cell walls, samples were removed and subjected to SDS-PAGE and Western blotting. As described previously (3.30), sample I comprised protein remaining in solution after interaction with and then removal of the wall by microcentrifugation. Sample IW consisted of protein released from the wall by boiling in the presence of sample buffer and sample II the total protein added to the system. Control assays, samples III and IIIW, were carried out in the absence of added α-amylase or HSA to confirm that the observed bands were specific for the proteins of interest.

When the binding of AmyL to cell walls was investigated, bands corresponding to AmyL were present in all samples (Fig. 8.4 A). Of the α-amylase added to the system (sample II) the amount of AmyL recovered from the wall (sample IW) was negligible (<10%). This was confirmed by the presence of considerable amounts of AmyL left in solution following wall binding and the removal of the wall (sample I).

In contrast, AmyLQS50.5 was found to bind to cell walls as indicated by the recovery of more AmyLQS50.5 (ca. 30%) by boiling in the presence of sample buffer (IW) and also the absence of AmyLQS50.5 after binding and removal of the wall (sample I).

When the wall binding of HSA was investigated, most of the protein was detected in sample I, suggesting that, as with AmyL, this eukaryotic protein did not bind significantly to isolated cell walls (Fig. 8.4 A). Some HSA was recovered from the walls after boiling with sample buffer (sample IW), but the amount was negligible.

Bands cross-reacting with anti-AmyL or anti-HSA antisera were not detected in the negative control samples (Fig. 8.4 B), although the molecular size standards were clearly visible in each case. This confirmed that the bands shown in Fig. 8.4 A were specific for the α-amylases and HSA.

The amphoteric nature of proteins dictates that, at pH values above or below their pI, proteins possess overall negative or positive charges, respectively. The chimeric α-amylase, AmyLQS50.5, was designed to have an overall positive charge and, consistent with the engineered changes, the pI value of this enzyme was found to be 10.0 (7.7). The IVWB assay was carried out at pH 7.2, a pH at which AmyLQS50.5 would be expected to be positively charged. In contrast, under the conditions of the assay, AmyL would be expected to have a slight negative charge since the pI of this protein was 7.0 (7.7). HSA has a measured pI of 5.2 (Denton and Harris, 1995) and would be expected to be negatively charged under the conditions of the IVWB assay.

Of the three proteins tested, only the most basic, AmyLQS50.5, showed significant binding to isolated cell walls. The relatively small amount of AmyL and HSA that was recovered from the walls by boiling was likely to have been carried over in the "void volume" of the wall matrix.

AmyL and AmyLQS50.5 had similar temperature and pH optima (7.2.1) indicating that their structural conformations were also similar, although the structural stability (7.2.2 and 7.3) and specific activity (7.6) were reduced. Therefore, under the IVWB assay conditions, the only major difference between the two α-amylases was likely to be their overall charge. This data suggests that, when in its native conformation, the overall charge of an exoprotein influences the degree to which the protein interacts with, and binds to, isolated cell walls *in vitro*. AmyL did not interact significantly with isolated cell walls, consistent with studies on the secretion of this α-amylase *in vivo* (8.1).

HSA did not appear to bind significantly to walls isolated from *B. subtilis* 168. This appears to contradict data obtained by Saunders *et al* (1987) who suggested that the failure to secrete HSA from walled *B. subtilis* was due to the binding of HSA to the cell wall at a stage following translocation. However, it is possible that exoproteins interact with the cell wall before the fully folded conformation is reached as a consequence of electrostatic interactions with basic amino acids which are not exposed when the protein is in its fully folded conformation.

In summary, the *in vitro* cell wall binding assay developed during this study provided a useful tool for assessing the ability of proteins in their native conformations to bind to cell walls. Although not considered here, the assay system will be used to study the effect of cell wall composition on the wall binding of various secretory proteins of prokaryotic or eukaryotic origin. The composition of the cell wall is likely to have a marked affect on the binding capacity of the peptidoglycan matrix and the anionic polymer content and degree of alanylation of the walls will be investigated specifically. Further, in parallel

with pulse-chase experiments, chemical cross-linking of proteins to cell walls during growth provides an *in vivo* approach to investigate the interaction of exoproteins with the *B. subtilis* cell wall and could be used to investigate the wall binding of AmyLQS50.5.

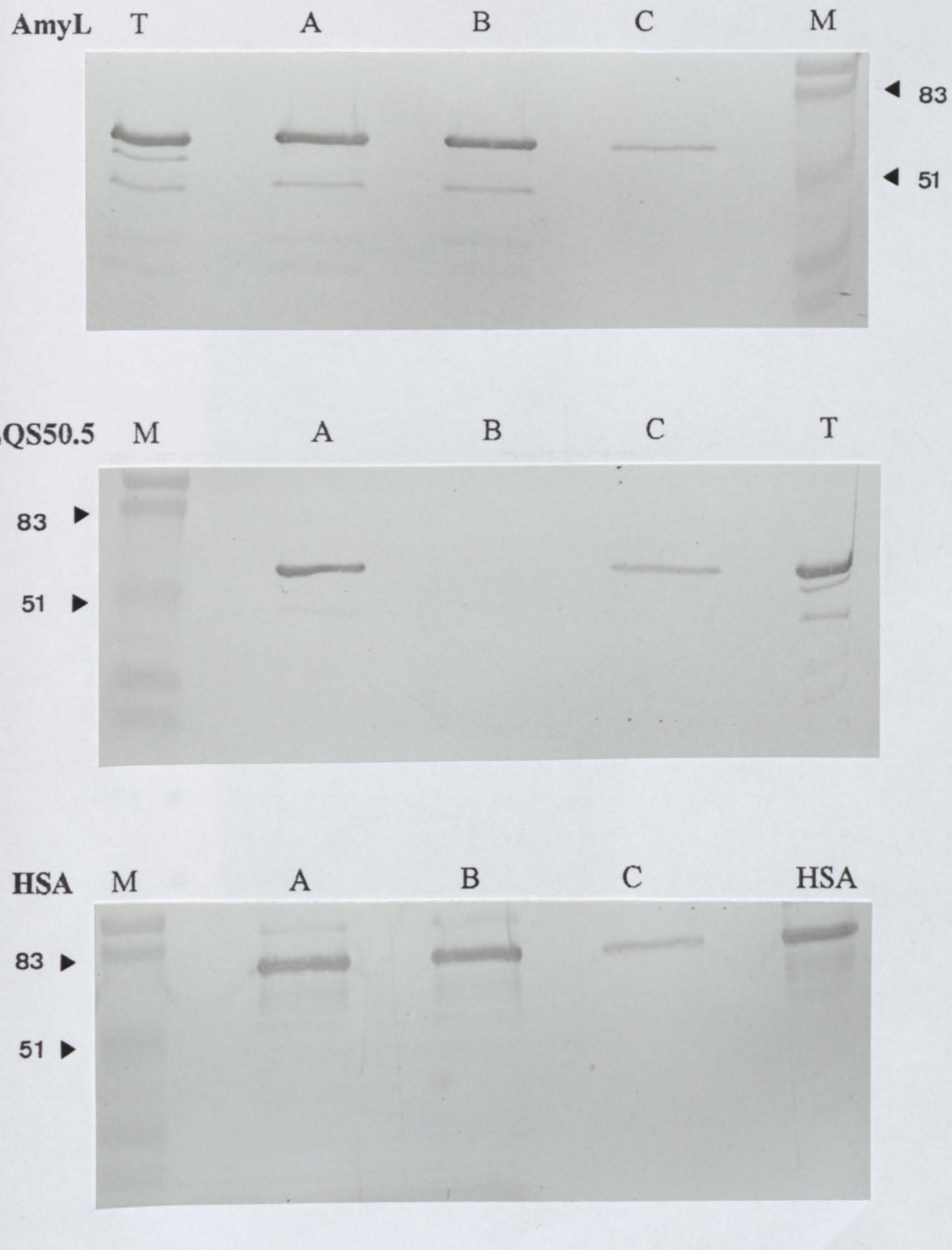
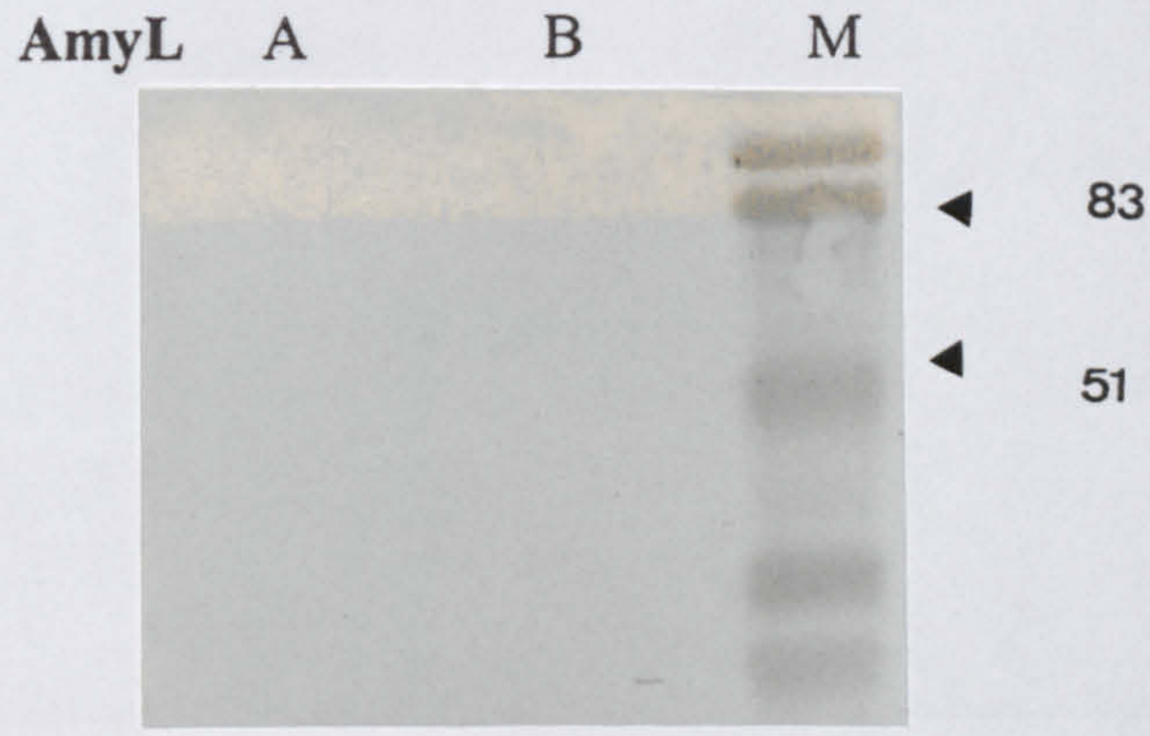


Fig. 8.4 A. IVWB assay Western blots.
Key:

T - Termamyl α-amylase standard
M - Pre-stained markers (kDa)
HSA - Albumin standard
A - Sample II (total protein)
B - Sample I (unbound protein)
C - Sample IW (wall bound protein)



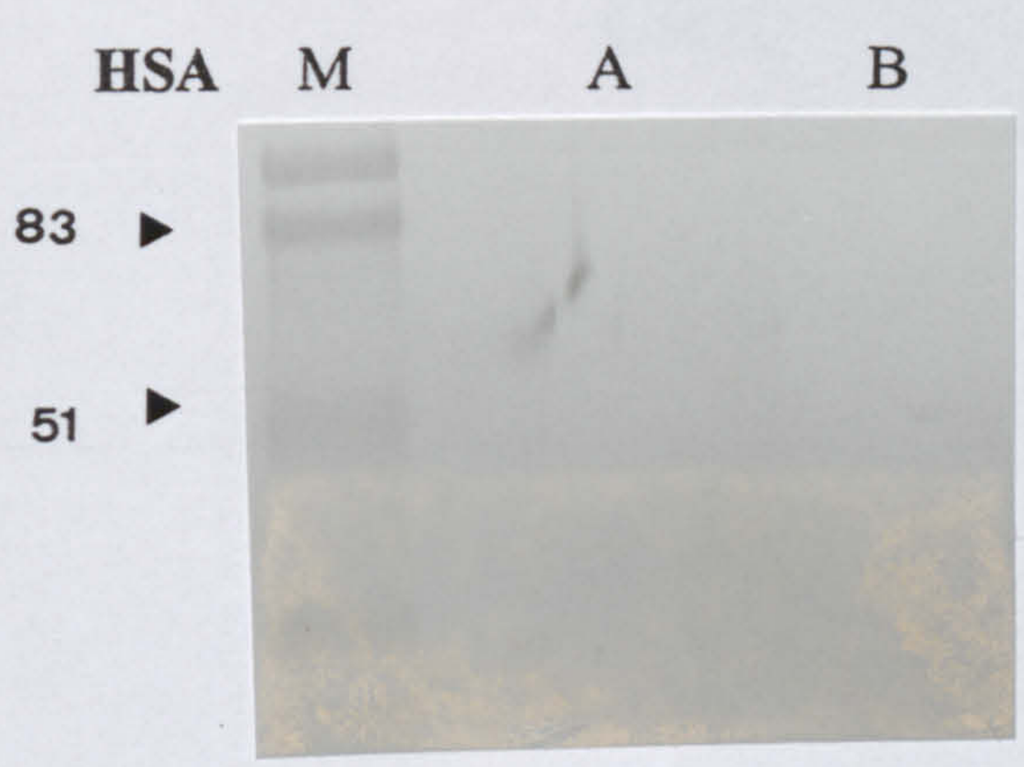


Fig. 8.4 B. IVWB assay negative control Western blots.
Key:
M - Pre-stained markers (kDa)
A - Sample III (unbound protein)
B - Sample IIIW (wall bound protein)

CHAPTER 9

GENERAL CONCLUSIONS

At the beginning of these studies there was little information on the potential influence of the cell wall on protein secretion from *B. subtilis*. In collaboration with Novo Nordisk we developed a novel approach to investigate the importance of the cell wall in secretion by engineering a normally well secreted *Bacillus* α-amylase, AmyL, in such a way as to promote charge interactions with the *B. subtilis* wall.

The PCR-based SOE technique was used to construct the component DNA fragments of a chimeric α-amylase gene, *amyLQS50*. SOE proved to be an effective tool for creating precise gene fusions in the absence of suitable restriction sites. The strategy allowed five specific regions of *amyL* to be replaced with corresponding regions of *amyQ* and *amyS*. In all cases, DNA sequencing verified that the gene splicing reactions had occurred as planned. However, nucleotide misincorporation was common in the generated DNA fragments, resulting probably from the use of a non-proof-reading *Taq* DNA polymerase. By assembling combinations of *amyLQS50* and *amyL* fragments, a range of chimeric α-amylase genes were constructed.

When shuttle plasmids encoding chimeric α-amylases were introduced into *B. subtilis* DN1885, detectable α-amylase activity was observed in some, but not all, strains. AmyLQS50.1, AmyLQS50.2 and AmyLQS50.5 were detected in culture supernatants on the basis of their enzymatic activity and Western blotting. Very low levels of AmyLQS50.3 were detected by Western blotting but enzymatic activity could not be detected in supernatants. Neither enzymatic activity or antibody cross-reactivity were observed with AmyLQS50 and AmyLQS50.4.

Compared to AmyL, the specific activity of purified AmyLQS50.5 was reduced by 2-fold and consequently, specific activity differences were also expected for the other chimeric α-amylases. The data suggests that the lack of detectable α-amylase activity in the supernatants of certain strains was the result of a combination of lower enzyme specific activities and reduced amounts of the α-amylases in the supernatants. Additionally, changes in the levels of gene expression can not be discounted at this stage. At the end of exponential growth the concentration of AmyLQS50.5 in the supernatant, as determined by enzymatic assay, was 4-fold lower than AmyL. Taking into account the differences in enzyme specific activities, this represented a 50% reduction in the amount of AmyLQS50.5 protein, in relation to AmyL.

Whenever the LQS1 fragment was included in a chimeric gene construct, α-amylase activity could not be detected. Replacement of fragment 1 from *amyLQS50* with WT1 lead to the construction of *amyLQS50.5*. The protein encoded by this gene exhibits α-amylase activity when expressed in *B. subtilis*.

The replacement of the LQS1 fragment of *amyLQS50* with WT1 resulted in the substitution of two AmyLQS50 amino acids, arginine-51 and serine-52, with AmyL amino acids glutamine-51 and alanine-52. This indicates that the substitution of just two amino acid residues had a marked affect on the quantity and quality of secreted α-amylase. The determination of the crystal structure of AmyL allowed the identification of a number of structurally and functionally important amino acid residues (Machius *et al.*, 1995) however, neither glutamine-51 or alanine-52 were implicated as being essential for enzymatic activity and/or structural integrity.

The *amyLQS50.1* and *amyLQS50.5* genes encoded catalytically active α-amylases which displayed considerably increased overall positive charge. When compared with the wild type, the temperature and pH optima of the most basic chimeric enzyme, AmyLQS50.5, were largely unaffected by the engineered changes. At high temperature or in the presence of the chelator EDTA, the structural stability of AmyLQS50.5 was adversely affected, resulting in rapid inactivation and/or degradation. AmyL was also affected by elevated temperature and the presence of EDTA, but the effects were not as pronounced as for AmyLQS50.5. Disruptions to the internal packing arrangements of AmyLQS50.5 caused by the presence of a large number of positively charged amino acid residues were expected and it is likely that this has contributed to the reduced structural stability of this enzyme.

The kinetics of processing and secretion of AmyLQS50.5 and AmyL were different. Both enzymes were processed rapidly but there was a 3-fold reduction in the rate of processing of AmyLQS50.5. A peak of mature AmyLQS50.5 occurred at 30 sec post-chase. An equivalent peak was not observed for AmyL and it was assumed that the AmyL peak occurred before the first sample was removed. Concurrent with or immediately following translocation, significant amounts of the mature forms of AmyLQS50.5 and AmyL were degraded. The amount of mature AmyL decreased initially and then remained constant, while AmyLQS50.5 appeared to be degraded completely. The rapid degradation of AmyLQS50.5 in exponentially growing cells lead to virtually no release of protein into the culture supernatant. AmyLQS50.5 is capable of being secreted and released, since this enzyme can be detected in the supernatants of long term (stationary phase) cultures. In contrast, during exponential growth, AmyL passed rapidly through the cell wall and was released unimpeded into the supernatant.

During secretion both AmyL and AmyLQS50.5 were subject to degradation, with the latter enzyme affected to a greater extent. The degradation of the wild type enzyme may reflect the fact that this enzyme is not native to *B. subtilis*. The protease(s) responsible for this "early" degradation is unknown

however, since the degradation occurs co- and/or post-translocationally, it is likely to be associated with the cytoplasmic membrane or cell wall. This would localize its activity in close proximity to the α-amylases as they are translocated. It is reasonable to assume that the protease responsible for the degradation is expressed in exponentially growing cells.

The majority of the extracellular proteases currently identified in *B. subtilis* are only synthesized after the end of exponential growth and consequently are unlikely to be responsible for the observed degradation of AmyLQS50.5 and AmyL. By virtue of their likely cellular locations there are two *B. subtilis* proteases identified to date which may be involved in the post-translocational degradation of AmyLQS50.5 and AmyL. The first, Epr, is a minor extracellular protease which has been putatively identified as being wall-bound (Sloma *et al.*, 1988; Pero and Sloma, 1993). The second candidate protease, FtsH, is an intracellular membrane-bound ATP-dependent protease (Tomoyasu *et al.*, 1993; Geisler and Schumann, 1993; Deuerling *et al.*, 1995). To demonstrate the role, if any, of these proteases in the degradation of the α-amylases, pulse-chase experiments will be carried out in *epr*- and *ftsH*-deficient backgrounds.

Levansucrase, which is produced by *B. subtilis* during exponential growth, is secreted by a two step mechanism. The first step is the insertion of the preprotein into the membrane and the removal of the N-terminal signal peptide. The second step, which is rate limiting, involves folding into its native conformation and release from the cytoplasmic membrane (Petit-Glatron *et al.*, 1993; Chambert *et al.*, 1995). This second step is modulated by ions such as Ca^{2+}, Fe^{3+} and H^{+} which promote rapid folding on the outside of the cytoplasmic membrane. The cell wall is able to concentrate Ca^{2+} ions at the cytoplasmic membrane/wall interface up to 100-fold higher than the concentration in the growth medium (Petit-Glatron *et al.*, 1993). Consequently, in addition to providing a buffer of the necessary ions for proper membrane functionality, the cell wall is involved in secretion by maintaining a concentration of divalent metal ions which promotes efficient exoprotein folding on the outside of the cytoplasmic membrane.

Compared with that of the wild type, the rate of *in vitro* refolding of AmyLQS50.5 was reduced by 3-fold, as determined by resistance to proteolytic degradation. On the basis of the data obtained in these studies, we have developed a working hypothesis which is applicable to the late stages of the secretion of AmyLQS50.5 and AmyL and is likely to be generally applicable to the secretion of other proteins from *B. subtilis*. We believe that the rapid and efficient folding of exoproteins on the outer surface of the cytoplasmic membrane is imperative for efficient secretion and exoproteins with reduced folding kinetics

are subjected to extensive proteolytic degradation before the native (fully folded) conformation is reached, resulting in significantly lower amounts of protein being released into the culture medium.

An obvious question is why should *B. subtilis* possess a protease which has the potential to severely limit the secretion of exoproteins? Since the second stage of protein secretion appears to be rate limiting, it is possible that the protease performs a clearing function to remove wrongly folded or slow folding exoproteins from the translocation complex. This ensures that the translocation complex does not become blocked since this has potentially lethal consequences for the cell. Native *B. subtilis* exoproteins are likely to have evolved to minimize degradation by this protease by ensuring that folding into a protease resistant conformation occurs rapidly.

The proper folding of *B. subtilis* exoproteins is assisted by the extracytoplasmic chaperone PrsA and simultaneous overexpression of exoenzymes and PrsA can increase the yield of released protein by up to 6-fold (Kontinen *et al.*, 1991; Jacobs *et al.*, 1993; Kontinen and Sarvas, 1993). The proposed peptidylprolyl-*cis*-*trans*-isomerase activity of PrsA promotes folding by catalyzing the isomerization of peptide bonds with a proline residue in the second position. Preliminary experiments on the secretion of AmyLQS50.5 (and other chimeric α-amylases) in the presence of the PrsA encoding plasmid, pKTH277 (provided by M Sarvas), did not increase the yield of released α-amylase, suggesting that PrsA had little or no affect on the secretion of AmyLQS50.5 expressed using the chromosome-based expression system. However, it is possible that this simply reflects the lower α-amylase copy number of the chromosomal expression system since PrsA only influences the secretion of exoproteins expressed at a high level. The presence of a propeptide can also mediate the proper folding of exoproteins (Cash *et al.*, 1989; Zhu *et al.*, 1989; Simonen and Palva, 1993). It is possible that "slower" folding *B. subtilis* exoproteins have to be equipped with a propeptide to increase the rate of folding at the cytoplasmic membrane/cell wall interface, thereby reducing proteolysis and ensuring efficient secretion.

Ca^{2+} ions are required for the activity, structural integrity and folding of α-amylases. In contrast to AmyL, preliminary data suggests that purified AmyLQS50.5 has a much reduced (~10-fold) binding affinity for Ca^{2+}, as estimated by the binding of $^{45}CaCl_2$. Three amino acids residues make up the calcium binding site of AmyL, asparagine-104, aspartate-200 and histidine-235. All three calcium binding residues are conserved in AmyLQS50.5. However, it is not known whether, in a manner analogous to levansucrase (Petit-Glatron *et al.*, 1993), AmyL possesses other, lower affinity calcium binding sites which were

affected by the engineered changes to this enzyme. Preliminary experiments revealed that the rate of refolding of AmyLQS50.5 and AmyL was affected by Ca^{2+}concentration, however, at all concentrations tested, the rate of refolding of AmyLQS50.5 was slower than AmyL, possibly as a consequence of its reduced calcium binding affinity.

Exoproteins can have one of three fates following folding into their native conformations. They can remain associated with the cytoplasmic membrane, associated with the cell wall or be released into the culture supernatant. It is assumed that proteins destined for release into the supernatant traverse the cell wall in the last stage of the export process. In general, rapidly exported proteins have a pI close to 7.0 and so have low net charge. AmyLQS50.5 was engineered to have a high pI and consequently, when in its native conformation, binds to cell walls isolated from *B. subtilis*. In contrast, AmyL does not bind significantly to cell walls on the basis of electrostatic interactions, confirming that the overall charge of an exoprotein can promote electrostatic interactions with the wall.

Because of the close proximity of the wall cylinder to the cytoplasmic membrane, it is possible that secretory proteins interact with the wall before they are fully folded. If such a situation does exist, slower folding proteins would be expected to be affected more than proteins with faster folding kinetics. When in its native conformation, HSA did not bind significantly to *B. subtilis* cell walls *in vitro*. It is possible that *in vivo* HSA binds to the wall before it is fully folded and it is conceivable that this wall-bound form never reaches the fully folded conformation or is degraded. This is supported by preliminary data on the *in vitro* refolding of AmyLQS50.5 in the presence and absence of isolated cell wall. The data suggests that the presence of the cell wall prevents a large proportion of the protein folding into an active form. In contrast, the refolding of AmyL was unaffected by the presence of wall.

The hypothesis that we have developed as a result of our work and that of others is summarized below in the form of a model (Fig. 9.1). Following translation, the exoproteins are translocated across the cytoplasmic membrane with the concomitant removal of the signal peptide. Folding occurs at the cytoplasmic membrane/cell wall interface and the rate at which the fully-folded conformation is reached determines the degree of post-translocational degradation. Even in the case of rapidly folding exoproteins, such as AmyL, some degradation occurs. Exoproteins with slower folding kinetics, such as AmyLQS50.5, are subjected to extensive degradation, by an as yet unidentified protease(s), whilst not in their fully folded state; the rate of degradation being inversely related to the rate of folding.

Key:

Anionic polymer

Ca^{2+} ion

Unidentified membrane- or wall-associated protease(s)

Unfolded/partially folded protein

Folded protein

Degraded protein

Translocation complex

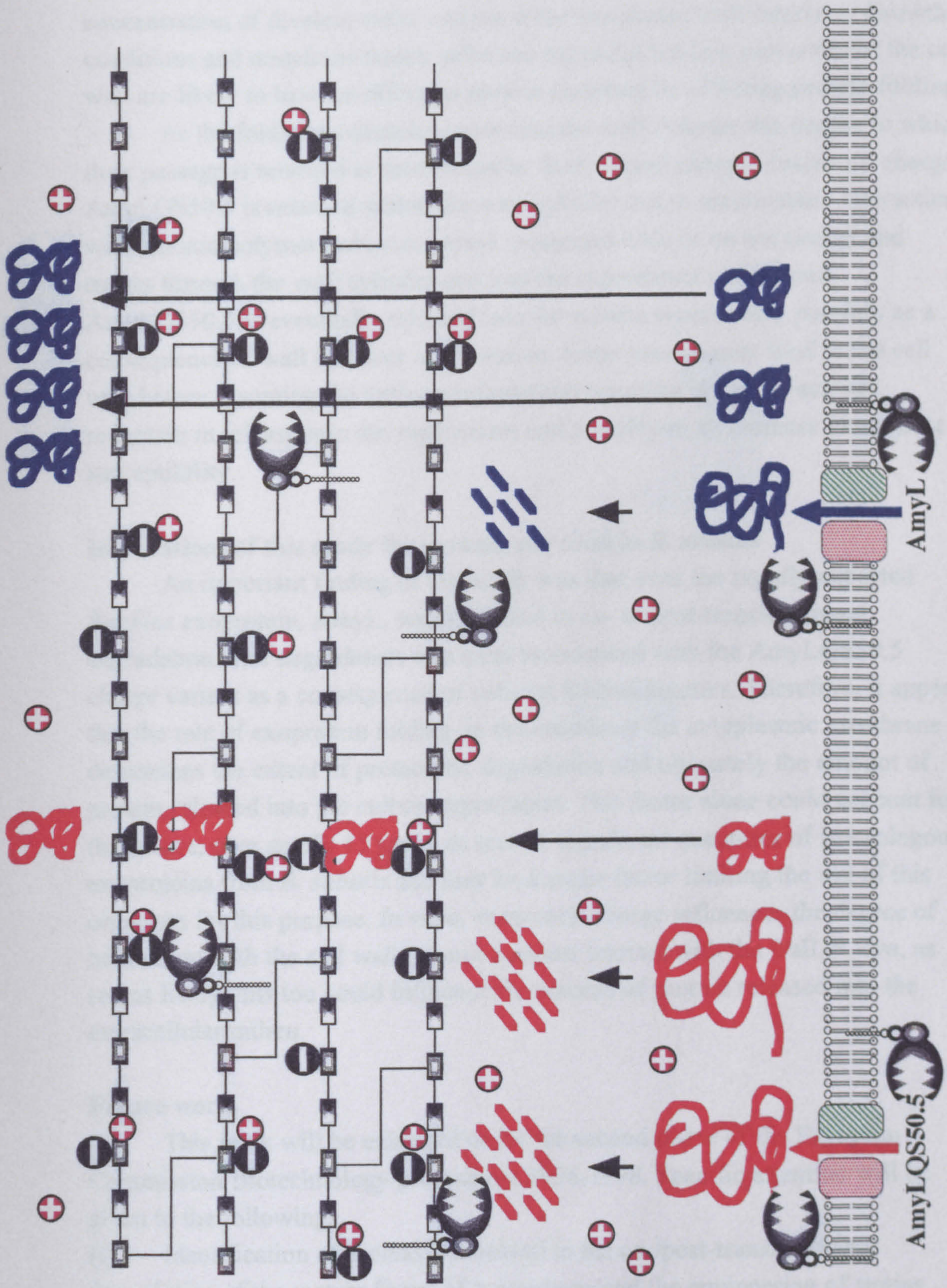


Fig. 9.1. The proposed model of α-amylase secretion.

The rate of folding is influenced by properties of individual exoproteins and also external factors such as the presence of chaperones and the concentration of divalent metal cations at the membrane/wall interface. Growth conditions and mutations which influence the metal binding properties of the cell wall are likely to have an effect on protein secretion by affecting protein folding.

As the folded exoproteins move into the wall cylinder the degree to which their passage is retarded is determined by their overall charge. Positively charged AmyLQS50.5 is retained within the wall cylinder due to electrostatic interactions with anionic polymers whereas, AmyL possesses little or no net charge and moves through the wall cylinder and into the supernatant unhindered. AmyLQS50.5 is eventually released into the culture supernatant, possibly as a consequence of wall turnover or saturation. Some protein may bind to the cell wall before assuming the native conformation resulting in loss of activity, reduction in release into the supernatant and possibly to an increase in protease susceptibility.

Implications of this study for protein secretion in *B. subtilis*

An important finding of this study was that even the rapidly exported *Bacillus* exoprotein, AmyL, was subjected to co- or post-translocational degradation. The degradation was most pronounced with the AmyLQS50.5 charge variant as a consequence of reduced folding kinetics. Therefore, it appears that the rate of exoprotein folding on the outside of the cytoplasmic membrane determines the extent of proteolytic degradation and ultimately the amount of protein released into the culture supernatant. This factor alone could account for the failure, over nearly 15 years, to secrete significant quantities of heterologous exoproteins from *B. subtilis* and may be a major factor limiting the use of this organism for this purpose. *In vitro*, exoprotein charge influences the degree of interaction with the cell wall. If such proteins interact with the wall *in vivo*, as seems likely, this too could influence the amount of protein released into the extracellular milieu.

Future work

This work will be extended under the second phase of the European Commission Biotechnology programme 1994-1998. Specific attention will be given to the following:

(i) Identification of proteases involved in the co-/post-translocational degradation of the mature forms of α-amylases and the engineering of strains optimized for native and heterologous protein export.

(ii) Analysis of the levels of chimeric α-amylase gene expression using Northern blotting.
(iii) The influence of Ca^{2+} concentration at the cytoplasmic membrane/cell wall interface on the folding of exoproteins.
(iv) The influence of cell wall composition on the *in vitro* binding and secretion of chimeric α-amylases.
(v) Investigation of the role of AmyL amino acid residues 51 and 52 on the secretability of α-amylase constructs.

CHAPTER 10

REFERENCES

Akimaru, J., Matsuyama, S., Tokuda, H. and Mizushima, S. (1991). Reconstitution of a protein translocation system containing purified SecY, SecE and SecA from *Escherichia coli*. Proceedings of National Academy of Sciences, USA **88**: 6545-6549.

Akiyama, Y. and Iko, K. (1987). Topology analysis of the SecY protein, an integral membrane protein involved in protein export in *Escherichia coli*. EMBO Journal **11**: 3465-3470.

Anderson, J. A., Green, R. S. and Archibald, A. R. (1978). Wall composition and phage binding properties of *Bacillus subtilis* W23 grown in chemostat culture media containing varied concentrations of phosphate. FEMS Microbiology Letters **4**: 129-132.

Archibald, A. R. (1976). Cell wall assembly in *Bacillus subtilis*: development of bacteriophage-binding properties as a result of the pulsed incorporation of teichoic acid. Journal of Bacteriology **127**: 956-960.

Archibald, A. R. (1985). Structure and assembly of the cell wall in *Bacillus subtilis*. Biochemal Society Transactions **13**: 990-993.

Archibald, A. R. and Coapes, H. E. (1976). Bacteriophage SP50 as a marker for cell wall growth in *Bacillus subtilis*. Journal of Bacteriology **125**: 1195-1206.

Archibald, A. R., Hancock, I. C. and Harwood, C. R. (1993). Cell wall structure, synthesis and turnover. In: *Bacillus subtilis* and other Gram-positive bacteria. Biochemistry, physiology, and molecular genetics (Sonenshein, A. L,. Hoch, J. A. and Losick, R. eds). American Society for Microbiology, Washington DC, USA.

Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H. Frank, G. and Tratschin, J. D. (1979). Thermal stability and protein structure. Biochemistry **18**: 5698-5703.

Beveridge, T. J. and Murray, R. G. E. (1976). Uptake and retention of metals by cell walls of *Bacillus subtilis*. Journal of Bacteriology **127**: 1505-1518.

Bieker, K. L. and Silhavy, T. J. (1990). PrlA (SecY) and PrlG (SecE) interact directly and function sequentially during protein translocation in *Escherichia coli*. Cell **61**: 833-842.

Bieker-Brady, K. and Silhavy, T. J. (1992). Suppressor analysis suggests a multistep, cyclic mechanism for protein secretion in *Escherichia coli*. EMBO Journal **11**: 3165-3174.

Bilgin, N., Lee, J. I., Zhu, H., Dalbey, R. and von Heijne, G. (1990). Mapping of catalytically important domains in *Escherichia coli* leader peptidase. EMBO Journal 9: 2717-2722.

Birdsell, D. C., Doyle, R. J. and Morgenstern, M. (1975). Organsization of teichoic acid in the cell wall of *Bacillus subtilis*. Journal of Bacteriology **121**: 726-734.

Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L. and Woldike, H. F. (1990). Calcium binding in α-amylases: an X-ray diffraction study at 2.1 Å resolution of two enzymes from *Aspergillus*. Biochemistry **29**: 6244-6249.

Boylan, R. J., Mendelson, N. H., Brooks, D. and Young, F. E. (1972). Regulation of the cell wall: analysis of a mutant of *Bacillus subtilis* defective in biosynthesis of teichoic acid. Journal of Bacteriology **110**: 281-290.

Bron, S. and Luxen, E. (1985). Segregational instability of pUB110-derived recombinant plasmids in *B.subtilis*. Plasmid. **14**: 235-244.

Bruckner, R., Shoseyov, O. and Doi, R. H. (1990). Multiple active forms of a novel serine protease from *Bacillus subtilis*. Molecular and General Genetics **221**: 486-490.

Brundage, L., Fimmel, C. J., Mizushima, S. and Wickner, W. (1992). SecY, SecE, and band 1 form the membrane-embedded domain of *Escherichia coli* preprotein translocase. Journal of Biological Chemistry **267**: 4166-4170.

Brundage, L., Hendrick, J. P., Scheibel, E., Driessen, A. J, M. and Wickner, W. (1990). The purified *Escherichia coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. Cell **62**: 649-657.

Buisson, G., Duee, E., Haser, R. and Payan, F. (1987). Three dimensional structure of porcine pancreatic α-amylase at 2.9 Å resolution. Role of calcium in structure and activity. EMBO Journal **6**: 3909-3916.

Canosi, U., Morelli, G. and Trautner, T. A. (1978). The relationship between molecular structure and transformation efficiency of some *Staphylococcus aureus* plasmids isolated from *Bacillus subtilis*. Molecular and General Genetics **166**: 259-267.

Carrascosa, J. L., Abella, G., Marco, S. and Carazo, J. M. (1990). Three-dimensioanl structure of the seven-fold form of *Bacillus subtilis* GroEL chaperonin. Journal of Structural Biology **104**: 2-8.

Cash, P. W., Zhu, X., Ohta, Y., Tsao, J., Lackland, H., Mateos-Nevado, M. D., Inouye, M., Stein, S., Jordan, F. and Tous, G. I. (1989). Synthesis of the propeptide of subtilisin BPN'. Peptide Research **2**: 292-296.

Chambert, R., Haddaoui, E. A. and Petit-Glatron, M.-F. (1995). *Bacillus subtilis* levansucrase: the efficiency of the second stage of secretion is modulated by external factors assisting folding. Microbiolgy **141**: 997-1005.

Chang, B. -Y., Chen, K. -Y., Wen, Y. -D. and Liao, C. -T. (1994). The response of a *Bacillus subtilis* temperature-sensitive *sigA* mutant to heat stress. Journal of Bacteriology **176**: 3102-3110.

Cheah, S. C., Hussey, H., Hancock, I. C. and Baddiley, J. (1982). Control of synthesis of wall teichoic acid during balanced growth of *Bacillus subtilis* W23. Journal of General Microbiology **128**: 593-599.

Cheung, H. -Y. and Freese, E. (1985). Monovalent cations enable cell wall turnover of the turnover-deficient *lyt-15* mutant of *Bacillus subtilis*. Journal of Bacteriology. **161**: 1222-1225.

Chothia, C. and Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. EMBO Journal **5**: 823-826.

Clarke-Sturman, A. J., Archibald, A. R., Hancock, I. C., Harwood, C. R., Merad, T. and Hobot, J. A. (1989). Cell wall assembly in *Bacillus subtilis*: partial conservation of polar wall material and the effect of conditions on the pattern of incorporation of new material at the polar caps. Journal of General Microbiology **135**: 657-665.

Coley, J., Tarelli, E., Archibald, A. R. and Baddiley, J. (1978). The linkage between teichoic acid and peptidoglycan in bacterial cell walls. FEBS Letters **88**: 1-9.

Coxon, R., D. (1990). Factors affecting protein export from *Bacillus subtilis*. PhD Thesis, University of Newcastle, UK.

Coxon, R. D., Harwood, C. R. and Archibald, A. R. (1991). Protein export during growth of *Bacillus subtilis*: the effect of extracellular protease deficiency. Letters in Applied Microbiology **12**: 91-94.

Dalbey, R. E. and Wickner, W. (1985). Leader peptidase catalyzes the release of exported proteins from the outer surface of the *Escherichia coli* plasma membrane. Journal of Biological Chemistry **260**: 15925-15931.

Dalbey, R. E. and Wickner, W. (1987). Leader peptidase of *Escherichia coli*: critical role of a small domain in membrane assembly. Science **235**: 783-787.

Declerck, N., Joyet, P., Gaillardin, C. and Masson, J.-M. (1990). Use of amber suppressors to investigate the thermal stability of *Bacillus licheniformis* α-amylase. Journal of Biological Chemistry **265**: 15481-15488.

Denton, K. A. and Harris, R. (1995). High-performance capillary electrophoretic separation of human serum albumin using a neutral coated capillary. Journal of chromatography **705**: 335-341.

Deuerling, E., Paeslack, B. and Schumann, W. (1995). The *fisH* gene of *Bacillus subtilis* is transiently induced after osmotic and temperature upshift. Journal of Bacteriology **177**: 4105-4112.

Diderichsen, B., Outtrup, H., Schulein, M. and Norman, B. E. (1987). Chimeric enzymes. Euopean Patent Application EP-87305781.

Diderichsen, B., Wedsted, U., Hedegaard, L., Jensen, B. R. and Sjoholm, C. (1990). Cloning of *aldB*, which encodes α-acetolactate decarboxylase, an exoenzyme from *Bacillus brevis*. Journal of Bacteriology 172: 4315-4321.

Dodd, I. B. and Egan, J. B. (1990). Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Research. 19: 5019-5026.

Doyle, R. J., McDannel, M. L., Helman, J. R. and Streips, U. N. (1975). Distribution of teichoic acid in the cell wall of *Bacillus subtilis*. Journal of Bacteriology 122: 152-158.

Driessen, A. J. M. (1990). Precursor protein translocation by the *Escherichia coli* translocase is directed by the proton motive-force. EMBO Journal 11: 847-853.

Ellwood, D. C. and Tempest, D. W. (1969). Control of teichoic acid and teichuronic acid biosynthesis in chemostat cultures of *Bacillus subtilis* var. *niger*. Biochemical Journal 111: 1-5.

Ellwood, D. C. and Tempest, D. W. (1972). Effects of environment on bacterial wall content and compostion. Advances in Microbial Physiology 7: 83-117.

Fan, D. P., Beckman, B. E. and Gardner-Eckstrom, H. L. (1975). Mode of cell synthesis in Gram-positive Bacilli. Journal of Bacteriology 123: 1157-1162.

Fischer, W. H., Koch, U., Rosel, P. and Fiedler, F. (1981a). Effect of alanine ester substitution and other structural features of lipoteichoic acids on their inhibitory activity against *Staphylococcus*. Journal of Biological Chemistry 255: 4557-4562.

Fischer, W., Rosel, P. and Koch, H. U. (1981b). Effect of alanine ester substitution and other structural features of lipoteichoic acids on their inhibitory activity against autolysins of *Staphyloccocus aureus*. Journal of Bacteriology 146: 467-475.

Flaherty, K. M., DeLuca-Flaherty, C. and McKay, D. B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. Nature **346**: 623-628.

Fogarty, W. M. (1983). Microbial amylases. In: Microbial Enzymes and Biotechnology (Fogarty, W. M. ed). Applied Science, London.

Forsberg, C. W., Wyrick, P. B., Ward, J. B. and Rogers, H. J. (1973). Effect of phosphate limitation on the morphology and wall composition of *Bacillus licheniformis* and its phosphoglucomutase deficient mutants. Journal of Bacteriology **113**: 969-984.

Foster, S. J. (1992). Analysis of the autolysins of *Bacillus subtilis* 168 during vegetative growth and differentiation by using renaturing polyacrylamide electrophoresis. Journal of Bacteriology **174**: 464-470.

Foster, S. J. (1993). Molecular analysis of three major wall-associated proteins of *Bacillus subtilis* 168: evidence for processing of the product of a gene encoding a 258 kDa precursor two-domain ligand-binding protein. Molecular Microbiology **8**: 299-310.

Garcia, P., Garcia, J. L., Garcia, E., Sanchez-Puelles, J. M. and Lopez, R. (1990). Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. Gene **86**: 81-88.

Gardel, C., Johnson, K., Jacq, A. and Beckwith, J. (1990). The *secD* locus of *Escherichia coli* codes for two membrane proteins required for protein export. EMBO Journal 9: 3209-3216.

Gartner, D., Degenkolb, J., Ripperger, J. A. E., Allmasnsberger, R. and Hillen, W. (1992). Regulation of the *Bacillus subtilis* W23 xylose utilization operon: interaction of the Xyl repressor with the *xyl* operator and the inducer xylose. Molecular and General Genetics. **232**: 415-422.

Gartner, D., Geissendorfer, M. and Hillen, W. (1988). Expression of the *Bacillus subtilis xyl* operon is repressed at the level of transcription and is induced by xylose. Journal of Bacteriology **170**: 3102-3109.

Geisler, U. and Schumann, W. (1993). Isolation of stress mutants of *Bacillus subtilis* by a novel genetic method. FEMS Microbiology Letters **108**: 251-254.

Geller, B. L. (1991). Energy requirements for protein translocation across the *Eschericia coli* inner membrane. Molecular Microbiology **5**: 2093-2098.

Glaser, L. and Burger, M. M. (1964). The synthesis of teichoic acids. III. Glycosylation of polyglycerolphosphate. Journal of Biological Chemistry **239**: 3187-3191.

Gould, A. R., May, B. K. and Elliott, W. H. (1975). Release of extracellular enzymes from *Bacillus amyloliquefaciens*. Journal of Bacteriology. **122**: 34-40.

Gray, G. L. Mainzer, S. E., Rey, M. W., Lamsa, M. H., Kindle, K. L., Carmona, C. and Requadt, C. (1986). Structural genes encoding the thermophilic α-amylases of *Bacillus stearothermophilus* and *Bacillus licheniformis*. Journal of Bacteriology **166**: 635-643.

Gryczan, T. J. (1982). Molecular cloning in *Bacillus subtilis*. In: The Molecular Biology of the Bacilli (Dubnau, D. ed). New York, Academic Press Inc.

Haandrikman, A. J., Kok, J., Laan, H., Soemitro, S., Ledeboer, A. M., Konings, W. N. and Venema, G. (1989). Identification of a gene required for maturation of an extracellular lactococcal serine proteinase. Journal of Bacteriology **171**: 2789-2794.

Haddaoui, E., Petit-Glatron, M.-F. and Chambert, R. (1995). Characterization of a new cell-bound α-amylase in *Bacillus subtilis* 168 Marburg that is only immunologically related to the exocellular α-amylase. Journal of Bacteriology **177**: 5148-5150.

Hancock, I. C. (1983). Activation and inactivation of synthesis of secondary wall polymers in *Bacillus subtilis* W23. Archives in Microbiology. **134**: 222-226.

Hancock, I. C. (1985). The regulation of synthesis of wall polymers and wall assembly in *Bacillus*. Biochemistry Society Transmissions **13**: 994-996.

Hancock, I. C. and Baddiley, J. (1985). Biosynthesis of the bacterial envelope polymers teichoic acid and teichuronic acid. In: The enzymes of biological macromolecules (Martonosi, A., N. ed). New York, Plenum Press.

Hartl, F. -U., Lecker, S., Schiebel, E., Hendrick, J. P. and Wickner, W. (1990). The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *Escherichia coli* plasma membrane. Cell **63**: 269-279.

Hastrup, S. (1988). Analysis of the *Bacillus subtilis* xylose regulon. In: Genetics and Biotechnology of Bacilli vol. 2. New York, Academic Press Inc.

He, X. S., Bruckner, R. and Doi, R. H. (1991). The protease genes of *Bacillus subtilis*. Research in Microbiology **142**: 797-803.

Hepstinall, S., Archibald, A. R. and Baddiley, J. (1970). Teichoic acids and membrane function in bacteria. Nature **225**: 519-521.

Herbold, D. R. and Glaser, L. (1975a). *Bacillus subtilis* N-acetylmuramic acid L-alanine amidase. Journal of Biological Chemsitry **250**: 1676-1680.

Herbold, D. R. and Glaser, L. (1975b). Interaction of N-acetylmuramic acid L-alanine amidase with cell wall polymers. Journal of Biological Chemistry **250**: 7231-7238.

Holm, L., Koivula, A. K., Lehtovaara, P. M., Hemminki, A. and Knowles, J. K. C. (1990). Random mutagensis used to probe the structure of *Bacillus stearothermophilus* α-amylase. Protein Engineering **3**: 181-191.

Holmes, D. S. and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Analytical Biochemistry **114**: 193-197.

Honda, K., Nakamura, K., Nishiguchi, M. and Yamane, K. (1993). Cloning and characterization of a *Bacillus subtilis* gene encoding a homologue of the 54-kilodalton subunit of mammalian signal recognition particle and *Escherichia coli* Ffh. Journal of Bacteriology **175**: 4885-4894.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77**: 61-68.

Hueck, C. J. and Hillen, W. (1995). Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the Gram-positive bacteria? Molecular Microbiology **15**: 395-401.

Hueck, C. J., Hillen, W. and Saier Jr, M. H. (1994). Analysis of a *cis*-acting sequence mediating catabolite repression in Gram-positive bacteria. Research in Microbiology **145**: 503-518.

Hughes, A. H., Hancock, I. C. and Baddiley, J. (1973). The function of teichoic acids in cation control in bacterial membranes. Biochemistry Journal **132**: 83-93.

Iwasaki, H., Shimada, A. and Ito, E. (1986). Comparative studies of lipoteichoic acids from several *Bacillus* strains. Journal of Bacteriology **167**: 508-516.

Jacob, S., Allmansberger, R., Gartner, D. and Hillen, W. (1991). Catabolite repression of the operon for xylose utilization from *Bacillus subtilis* W23 is mediated at the level of transcription and depends on a *cis* site in the *xylA* reading frame. Molecular and General Genetics. **229**: 189-196.

Jacobs, M., Andersen, J. B., Kontinen, V. and Sarvas, M. (1993). *Bacilus subtilis* PrsA is required *in vivo* as an extracytoplasmic chaperone for secretion of active enzymes synthesized either with or without pro-sequences. Molecular Microbiology **8**: 957-966.

Janniere, L., Niaudet, B., Pierre, E. and Ehrlich, S. D. (1985). Stable gene amplification in the chromosome of *Bacillus subtilis*. Gene **40**: 47-55.

Jeong, S. M., Yoshikawa, H. and Takahashi, H. (1993). Isolation and characterization of the *secE* homologue gene of *Bacillus subtilis*. Molecular Microbiology **10**: 133-142.

Johnson, C., Chandrasekher, G. N. and Georgopoulos, C. (1989). *Escherichia coli* DnaK and GrpE heat shock proteins interact both *in vivo* and *in vitro*. Journal of Bacteriology **171**: 1590-1596.

Jolliffe, L. K, Doyle, R. J and Streips, U. N (1980). Extracellular proteases modify cell wall turnover in *Bacillus subtilis*. Journal of Bacteriology **141**: 1199-1208.

Jolliffe, L. K., Doyle, R. J. and Streips, U. N. (1981). The energized membrane and cellular autolysis in *Bacillus subtilis*. Cell **25**: 753-763.

Jørgensen, P. L., Hansen, C. K., Poulsen, G. B. and Diderichsen, B. (1990). *In vivo* genetic engineering: homologous recombination as a tool for plasmid construction. Gene **96**: 37-41.

Jørgensen, P. L., Poulsen, G. B. and Diderichsen, B. (1991a). Cloning of a chromosomal α-amylase gene from *Bacillus stearothermophilus*. FEMS Letters **77**: 271-276.

Jørgensen, S. T., Skov, K. W. and Diderichsen, B. (1991b). Cloning, sequence, and expression of a lipase gene from *Pseudomonas cepacia*: lipase production in heterologous hosts requires two *Pseudomonas* genes. Journal of Bacteriology **173**: 559-567.

Kamitani, S., Akiyama, Y. and Ito, K. (1992). Identification and characterization of an *Escherichia coli* gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme. EMBO Journal **11**: 57-62.

Kessler, R. E., Van de Rijn, I. and McCarty, M. (1979). Characterization and localization of the enzymatic deacylation of lipoteichoic acid in group A streptococci. Journal of Experimental Medicine **150**: 1498-1509.

Kieser, T. (1984). Factors affecting the isolation of CCC DNA from *Steptomyces lividans* and *Escherichia coli*. Plasmid **12**: 19-36.

Killian, J. A., de Jong, A. M. P., Bijvelt, B., Verkieij, A. J. and de Kruijff, B. (1990a). Induction of non-bilayer lipid structures by functional signal peptides. EMBO Journal 9: 815-819.

Killian, J. A., Keller, R. C A., Struyve, M., de Kroon, A. I. P. M., Tommassen, J. and de Kruijff, B. (1990b). Tryptophan fluorescence study on the interaction of the signal peptide of the *Escherichia coli* outer membrane protein PhoE with model membranes. Biochemistry **29**: 8131-8137.

Kim, I.-C., Cha, J.-H., Kim, J.-R., Jang, S.-Y., Seo, B.-C., Cheong, T.-K., Lee, D. S., Choi, Y. D. and Park, K. H. (1992). Catalytic properties of the cloned amylase from *Bacillus licheniformis*. Journal of Biological Chemistry **267**: 22108-22114.

Koch, A. L. (1983). The surface stress theory of microbial morphogenesis. Advances in Microbial Physiology **24**: 301-366.

Koch, A. L., Kirchner, G., Doyle, R. J. and Burdett, I. D. J. (1985). How does a *Bacillus* split its septum right down the middle? Annales de l'Institute Pasteur Microbiology **136**: 91-98.

Koch, H., U. and Fischer, W. (1978). Acyldiglucosyliacylglycerol-containing lipoteichoic acid with a poly (3-O-galabiosyl 2-O galactosyl-*sn*-glycero-1-phosphate) chain from *Streptococcus lactis* Kiel 42172. Biochemistry **17**: 5275-5281.

Kontinen, V. P., Saris, P. and Sarvas, M. (1991). A gene (*prsA*) of *Bacillus subtilis* involved in a novel, late stage of protein export. Molecular Microbiology. **5**: 1273-1283.

Kontinen, V. P. and Sarvas, M. (1988). Mutants of *Bacillus subtilis* defective in protein export. Journal of General Microbiology. **134**: 2333-2344.

Kontinen, V. P. and Sarvas, M. (1993). The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limlit for high level secretion. Molecular Microbiology **8**: 727-737.

Kraus, A., Hueck, C., Gartner, D. and Hillen, W. (1994). Catabolite repression of the *Bacillus subtilis xyl* operon involves a *cis* element functional in the context of an unrelated sequence, and glucose exerts an additional *xylR*-dependent repression. Journal of Bacteriology **176**: 1738-1745.

Kreuzer, P., Gartner, D., Allmansberger, R. and Hillen, W. (1989). Identification and sequence analysis of the *Bacillus subtilis* W23 *xylR* gene and *xyl* operator. Journal of Bacteriology. **172**: 3840-3845.

Krieg, U. C., Walter, P. and Johnson, A. E. (1986). Photo-crosslinking of the signal sequence of nascent preprolactin to the 54-kDa polypeptide of the signal recognition particle. Proceedings of National Academy of Sciences, USA **83**: 8604-8608.

Kumamoto, C. A. (1991). Molecular chaperones and protein translocation across the *Escherichia coli* inner membrane. Molecular Microbiology **5**: 19-22.

Kumamoto, C. A., Chen, L., Fandl, J. and Tai, P. C. (1989). Purification of the *Eschericia coli secB* gene product and demonstration of its activity in an *in vitro* protein translocation system. Journal of Biological Chemistry **264**: 2242-2249.

Kumamoto, C. A. (1989). *Escherichia coli* SecB associates with exported protein precursors *in-vivo*. Proceedings of National Academy of Sciences, USA **86**: 5320-5324.

Kuroda, A., Rashid, M. H. and Sekiguchi, J. (1992). Molecular cloning and sequencing of the upstream region of the major *Bacillus subtilis* autolysin gene: a modifier protein exhibiting sequence homology to the major autolysin *spoIID* product. Journal of General Microbiology **138**: 1067-1076.

Kuroda, A. and Sekiguchi, J. (1991). Molecular cloning and sequencing of a major *Bacillus subtilis* autolysin gene. Journal of Bacteriology **173**: 7304-7312.

Kurzchalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkaveva, E. S., Bielka, H. and Rapoport, T. A. (1986). The signal sequence of the nascent preprolactin interacts with the 54-kDa polypeptide of the signal recognition particle. Nature **320**: 634-636.

Kusukawa, N., Yura, T., Ueguchi, C., Akiyama, T. and Ito, K. (1989). Effects of mutations in heat-shock genes *groES* and *groEL* on protein export in *Escherichia coli*. EMBO Journal **8**: 3517-3521.

Laemmli, U. K. (1970). Cleavage of structural protein during the assembly of the head of bacteriophage T4. Nature **227**: 680-685.

Lang, W. K., Glassey, K. and Archibald, A. R. (1982). Influence of phosphate supply on teichoic acid and teichuronic acid content of *Bacillus subtilis* walls. Journal of Bacteriology **151**: 367-375.

Laoide, B. M., Chambliss, G. H. and McConnell, D. J. (1989). *Bacillus licheniformis* α-amylase gene, *amyL*, is subject to promoter-independent catabolite repression in *Bacillus subtilis*. Journal of Bacteriology **171**: 2435-2442.

Lazarevic, V., Margot, P., Soldo, B. and Karamata, D. (1992). Sequencing and analysis of the *Bacillus subtilis lytRABC* divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier. Journal of General Microbiology **138**: 1949-1961.

Lecker, S., Lill, R., Ziegelhoffer, T., Georgopoulos, C., Bassford, P. J., Kumamoto, C. A. and Wickner, W. (1989). Three pure chaperone proteins of *Escherichia coli* - SecB, trigger factor and GroEL - form soluble complexes with precursor proteins *in vitro*. EMBO Journal **8**: 2703-2709.

Li, M. and Wong, S.-L. (1992). Cloning and characterization of the *groESL* operon from *Bacillus subtilis*. Journal of Bacteriology **174**: 3981-3992.

Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. and Zylicz, M. (1991). *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. Proceedings of National Academy of Sciences, USA **88**: 2874-2878.

Lill, R., Cunningham, K., Brundage, A., Ito, K., Oliver, D. and Wickner, W. (1989). SecA protein hydrolyzes ATP and is an essential component of the protein translocase of *Escherichia coli*. EMBO Journal **8**: 961-966.

Lill, R., Dowhan, W. and Wickner, W. (1990). The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. Cell **60**: 271-280.

Liu, J. and Walsh, C. T. (1990). Peptidyl-prolyl *cis-trans*-isomerase from *Escherichia coli*: a periplasmic homologue of cyclophilin that is not inhibited by cyclosporin A. Proceedings of National Academy of Sciences, USA **87**: 4028-4032.

Luirink, J., High, S., Wood, H., Giner, A., Tollervey, D. and Dobberstein, B. (1992). Signal-sequence recognition by an *Escherichia coli* ribonucleoprotein complex. Nature **359**: 741-743.

Machius, M., Weigand, G. and Huber, R. (1995). Crystal structure of calcium-depleted *Bacillus licheniformis* α-amylase at 2.2 Å resolution. Journal of Molecular Biology **246**: 545-559.

Margot, P. and Karamata, D. (1992). Identification of the structural genes for N-acetylmuramoyl-L-alanine amidase and its modifier in *Bacillus subtilis* 168: inactivation of these genes by insertional mutagenesis has no effect on growth or cell separation. Molecular and General Genetics **232**: 359-366.

Margot, P., Mauel, C. and Karamata, D. (1991). The *Bacillus subtilis* N-acetylglucosaminidase is encoded by a monocistronic operon controlled by a σ^D dependent promoter. 6th International conference on Bacilli, Stanford, USA., Stanford University Press.

Matsuura, Y., Kusunoki, M., Harada, W. and Kakudo, M. (1984). Structure and possible catalytic residues of Taka-amylase A. Journal of Biochemistry **95**: 697-702.

Mauck, J., Chin, L. and Glaser, L. (1971). Turnover of cell wall of Gram-positive bacteria. Journal of Biological Chemistry **246**: 1820-1827.

Mauck, J. and Glaser, L. (1972). On the mode of *in vivo* assembly of the cell wall of *Bacillus subtilis*. Journal of Biological Chemistry **247**: 1810-1827.

Mauel, C., Young, M., Margot, P. and Karamata, D. (1989). The essential nature of teichoic acids in *Bacillus subtilis* as revealed by insertional mutagenesis. Molecular and General Genetics **215**: 388-394.

Merad, T., Archibald, A. R., Hancock, I. C., Harwood, C. R. and Hobot, J. A. (1989). Cell wall assembly in *Bacillus subtilis*: visualization of the old and new material by electron microscopic examination of samples stained selectively for teichoic acid and teichuronic acid. Journal of General Microbiology **135**: 645-655.

Meyer, D. I., Krause, E. and Dobberstein, B. (1982). Secretory protein translocation across membranes - the role of the "docking protein". Nature **297**: 647-650.

Miyakawa, Y. and Komano, T. (1980). Study of the cell cycle of *Bacillus subtilis* using temperature sensitive mutants, I. Isolation and genetic analysis of the mutants defective in septum formation. Molecular and General Genetics **181**: 207-214.

Mobley, H. L, T., Koch, A. L., Doyle, R. J. and Streips, U. N. (1984). Insertion and fate of the cell wall in *Bacillus subtilis*. Journal of Bacteriology **158**: 169-179.

Mullis, K. B. and Faloona, F., A. (1987). Specific synthesis of DNA *in vitro via* a polymerase-catalyzed chain reaction. Methods in Enzymology **155**: 335-350.

Nakamura, K., Nakamura, A., Takamatsu, H., Yoshikawa, H. and Yamane, K. (1990). Cloning and characterization of a gene homologous to the *Escherichia coli secY*. Journal of Biochemistry. **107**: 603-607.

Narberhaus, F. and Bahl, H. (1992). Cloning, sequencing and molecular analysis of the *groESL* operon of *Clostridium acetobutylicum*. Journal of Bacteriology **174**: 3282-3289.

Narberhaus, F., Giebeler, K. and Bahl, H. (1992). Molecular characterization of the *dnaK* gene region of *Clostridium acetobutylicum*, including *grpE*, *dnaJ*, and a new heat shock gene. Journal of Bacteriology **174**: 3290-3299.

Nishiyama, K., Mizushima, S. and Tokuda, H. (1993). A novel membrane protein involved in protein translocation across the cytoplasmic membrane of *Escherichia coli*. EMBO Journal **12**: 3409-3415.

Nishiyama, K.-I., Hanada, M. and Tokuda, H. (1994). Disruption of the gene encoding p12(SecG) reveals the direct involvement and important function of SecG in the protein translocation of *Escherichia coli* at low temperature. EMBO Journal **13**: 3272-3277.

Ohta, Y., Hojo, H., Aimoto, S., Kobayashi, T., Zhu, X., Jordan, F. and Inouye, M. (1991). Pro-peptide as an intermolecular chaperone: renaturation of denatured subtilisin E with a synthetic propeptide. Molecular Microbiology **5**: 1507-1510.

Ortlepp, S. A., Ollington, J. F. and McConnell, D. J. (1983). Molecular cloning in *Bacillus subtilis* of a *Bacillus licheniformis* gene encoding a thermostable alpha-amylase. Gene **23**: 267-276.

Park, S. S., Wong, S.-L., Wang, L. F. and Doi, R., H. (1989). *Bacillus subtilis* subtilisin gene (*aprE*) is expressed from a sig-A (sig-43) promoter *in vitro* and *in vivo*. Journal of Bacteriology **171**: 2657-2665.

Pero, J. and Sloma, A. (1993). Proteases. In: *Bacillus subtilis* and other Gram-positive bacteria (Sonenshein, A. L,. Hoch, J. A. and Losick, R. eds). Biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington DC, USA.

Perutz, M. F. (1978). Electrostatic effects in proteins. Science **201**: 1187-1191.

Petit-Glatron, M-F., Grajcar, L., Munz, A. and Chambert, R. (1993). The contribution of the cell wall to a transmembrane calcium gradient could play a key role in *Bacillus subtilis* protein secretion. Molecular Microbiology 9: 1097-1106.

Phillips, G. J. and Silhavy, T. J. (1990). Heat shock proteins DnaK and GroEL facilitate export of LacZ hybrid proteins in *Escherichia coli*. Nature **344**: 882-884.

Phillips, G. J. and Silhavy, T. J. (1992). The *Escherichia coli ffh* gene is necessary for viability and efficient protein export. Nature **359**: 744-746.

Power, S. D., Adams, R. M. and Wells, J. A. (1986). Secretion and autoproteolytic maturation of subtilisin. Proceedings of National Academy of Sciences, USA **83**: 3096-3100.

Priest, F. G. (1977). Extacellular enzyme synthesis in the genus *Bacillus*. Bacteriological Reviews. **41**: 711-753.

Pugsley, A. J. (1993). The complete general secretory pathway in Gram-negative bacteria. Microbiological Reviews **57**: 50-108.

Ribes, V., Romisch, K., Giner, A., Dobberstein, B. and Tollervey, D. (1990). *Escherichia coli* 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. Cell **63**: 591-600.

Robson, R. L. and Baddiley, J. (1977a). Morphological changes associated with novobiocin resistance in *Bacillus licheniformis*. Journal of Bacteriology **129**: 1045-1050.

Robson, R. L. and Baddiley, J. (1977b). Role of teichuronic acid in *Bacillus licheniformis*: defective autolysis due to deficiency of teichuronic acid in a novobiocin resistant mutant. Journal of Bacteriology **129**: 1051-1058.

Rogers, H. J., Perkins, H. R. and Ward, J. B. (1980). Microbial cell walls and membranes. Chapman and Hall, London.

Roitsch, C. A. and Hageman, J. H. (1983). Bacillopeptidase F: two forms of a glycoprotein serine protease from *Bacillus subtilis*. Journal of Bacteriology **155**: 145-152.

Romisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M. and Dobberstein, B. (1989). Homology of the 54K protein of signal recognition particle, docking protein and two *Escherichia coli* proteins with putative GTP-binding domains. Nature **340**: 478-482.

Rufo, G. A., Sullivan, B. J., Sloma, A. and Pero, J. (1990). Isolation and characterization of a novel extracellular metalloprotease from *Bacillus subtilis*. Journal of Bacteriology **172**: 1019-1023.

Sadaie, Y. (1989). Molecular cloning of a *Bacillus subtilis* gene involved in cell division, sporulation and exoenzyme secretion. Japanase Journal of Genetics **64**: 111-119.

Sadaie, Y. and Kada, T. (1983). Effect of septum initiation mutation on sporulation and competent cell formation in *Bacillus subtilis*. Molecular and General Genetics **190**: 176-178.

Sadaie, Y., Takamatsu, H., Nakamura, K. and Yamane, K. (1991). Sequencing reveals similarity of the wild-type *div*$^+$ gene of *Bacillus subtilis* to the *Escherichia coli secA* gene. Gene **98**: 101-105.

Sadaie, Y. and Kada, T. (1985). *Bacillus subtilis* gene involved in cell division, sporulation and exoenzyme secretion. Journal of Bateriology **163**: 648-653.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of National Academy of Sciences, USA **74**: 5463-5467.

Saunders, C. W., Schmidt, B. J., Mallonee, R. L. and Guyer, M., S. (1987). Secretion of human serum albumin from *Bacillus subtilis*. Journal of Bacteriology. **169**: 2917-2925.

Schatz, P. J., Bieker, K. L., Ottemann, K. M., Silhavy, T. J. and Beckwith, J. (1991). One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the *Escherichia coli* secretion machinery. EMBO Journal **10**: 1749-1757.

Schatz, P. J., Riggs, P. D. and Beckwith, J. (1989). The *secE* gene encodes an integral membrane protein required for protein export in *Escherichia coli*. Genes and Development **3**: 1035-1044.

Scherrer, R. and Gerhardt, P. (1971). Molecular sieving by the *Bacillus megaterium* cell wall and protoplast. Journal of Bacteriology **107**: 718-735.

Schmidt, A., Schiesswohl, M., Volker, U., Hecker, M. and Schumann, W. (1992). Cloning, sequencing, mapping, and transcriptional analysis of the *groESL* operon from *Bacillus subtilis*. Journal of Bacteriology **174**: 3993-3999.

Schulz, A., Tzschaschel, B. and Schumann, W. (1995). Isolation and analysis of mutants of the *dnaK* operon of *Bacillus subtilis*. Molecular Microbiology **15**: 421-429.

Schulz, A.and Schumann, W. (1996). *hrcA*, the first gene of the *Bacillus subtilis dnaK* operon encodes a negative regulator of class I heat shock genes. Journal of Bacteriology **178**: 1088-1093.

Shabarova, Z. A., Hughes, N. A. and Baddiley, J. (1962). The influence of adjacent phosphate and hydroxyl groups on amino acid esters. Biochem. Journal **83**: 216-219.

Shiba, K., Ito, K., Yura, T. and Cerretti, D. P. (1984). A defined mutation in the protein export gene within the *spc* ribosomal operon of *Escherichia coli*: isolation and characterization of a new temperature-sensitive *secY* mutant. EMBO Journal **3**: 631-635.

Siegel, V. and Walter, P. (1988). Binding sites of the 19-kDa and 68/72-kDa signal recognition particle (SRP) proteins on SRP RNA as determined by protein-RNA "footprinting". Proceedings of National Academy of Sciences, USA **85**: 1801-1805.

Simon, S. M. and Blobel, G. (1992). Signal peptides open protein-conducting channels in *Escherichia coli*. Cell **69**: 677-684.

Simonen, M. and Palva, I. (1993). Protein secretion in *Bacillus* species. Microbiological Reviews. **57**: 109-137.

Sloma, A., Ally, A., Ally, D. and Pero, J. (1988). Gene encoding a minor extracellular protease in *Bacillus subtilis*. Journal of Bacteriology **170**: 5557-5563.

Sloma, A., Rudolph, C. F., Rufo, G. A., Sullivan, B. J., Theriault, K. A., Ally, D. and Pero, J. (1990a). Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*. Journal of Bacteriology **172**: 1024-1029.

Sloma, A., Rufo, G. A., Rudolph, C. F., Sullivan, B. J., Theriault, K. A. and Pero, J. (1990b). Bacillopeptidase F of *Bacillus subtilis*: purification and cloning of the gene. Journal of Bacteriology **172**: 1470-1477.

Sloma, A., Rufo, G. A., Theriault, K. A., Dwyer, M., Wilson, S. W. and Pero, J. (1991). Cloning and characterization of the gene for an additional extracellular serine protease from *Bacillus subtilis*. Journal of Bacteriology **173**: 6889-6895.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology **98**: 503-517.

Stahl, M. L. and Ferrari, E. (1984). Replacement of the *Bacillus subtilis* subtilisin structural gene with an *in vitro* derived deletion mutation. Journal of Bacteriology **158**: 411-418.

Stein, E. A. and Fischer, E. H. (1958). The resistance of α-amylases to proteolytic attack. Journal of Biological Chemistry **232**: 867-879.

Stephens, M. A., Ortlepp, S. A., Ollington, J. F. and McConnel, D. J. (1984). Nucleotide sequence of the 5' region of the *Bacillus licheniformis* α-amylase gene: comparison with the *Bacillus amyloliquefaciens* gene. Journal of Bacteriology **158**: 369-372.

Stewart, G. C. (1993). Catablolite repression the Gram-positive bacteria: generation of negative terminators of transcription. Journal of Cellular Biochemistry. **51**: 25-28.

Strauch, M. A., Spiegelman, G. B., Perego, M., Johnson, W. C., Burbulys, D. and Hoch, J. A. (1989). The transition state transcription regulator *abrB* of *Bacillus subtilis* is a DNA binding protein. EMBO Journal **8**: 1615-1621.

Studer, R. E. and Karamata, D. (1988). Cell wall proteins in *Bacillus subtilis*. In: Antibiotic inhibition of bacterial cell surface assembly and function. American Society for Microbiology, Washington DC, USA.

Sturman, A. J. and Archibald, A. R. (1978). Conservation of phage receptor material at the polar caps of *Bacillus subtilis* W23. FEMS Microbiology Letters **4**: 255-259.

Suh, J.-W., Boylan, S. A., Thomas, S. M., Dolan, K. M., Oliver, D. B. and Price, C. W. (1990). Isolation of a *secY* homologue from *Bacillus subtilis*: evidence for a common export pathway in eubacteria. Molecular Microbiology. **4**: 305-314.

Suzuki, Y., Ito, N., Yuuki, T., Yamagata, H. and Udaka, S. (1989). Amino acid residues stabilizing a *Bacillus* α-amylase against irreversible thermoinactivation. Journal of Biological Chemistry **264**: 18933-18938.

Takagi, M. and Imanaka, T. (1989). Role of the pre-pro region of neutral protease in secretion in *Bacillus subtilis*. Journal of Fermentation and Bioengineering **67**: 71-76.

Takagi, M., Imanaka, T. and Aiba, S. (1985). Nucleotide sequence and promoter region for the neutral protease gene *Bacillus stearothermophilus*. Journal of Bacteriology **163**: 824-831.

Takamatsu, H., Fuma, S.-I., Nakamura, K., Sadaie, Y., Shinkai, A., Matsuyama, S.-I., Mizushima, S. and Yamane, K. (1992). *In vivo* and *in vitro* characterization of the *secA* gene product of *Bacillus subtilis*. Journal of Bacteriology **174**: 4308-4316.

Takase, K., Mizuno, H. and Yamane, K. (1988). NH_2-terminal processing of *Bacillus subtilis* α-amylase. Journal of Biological Chemistry **263**: 11548-11553.

Takkinen, K., Pettersson, R. F., Kalkkinen, N., Palva, I., Soderlund, H. and Kaariainen, L. (1983). Amino acid sequence of α-amylase from *Bacillus amyloliqufaciens* deduced from the nucleotide sequence of the cloned gene. Journal of Biological Chemistry **258**: 1007-1013.

Tjalsma, H., Bron, S., Venema, G. and van Dijl, J. M. (1995). Type I signal peptidases of *Bacillus subtilis*. Abstract of the European *Bacillus* Protein Secretion Group conference, Helsinki, Finland.

Toma, S., Del Bue, M., Pirola, A. and Grandi, G. (1986). *npR1* and *npR2* regulatory regions for neutral protease expression in *Bacillus subtilis*. Journal of Bacteriology **167**: 740-743.

Tomazic, S. J. and Klibanov, A. M. (1988). Why is one *Bacillus* α-amylase more resistant against irreversible thermoinactivation than another? Journal of Biological Chemistry **263**: 3092-3096.

Tomoyasu, T., Yuki, T., Morimura, S., Mori, H., Yamanaka, K., Niki, H., Hiraga, S. and Ogura, T. (1993). The *Escherichia coli* FtsH protein is a prokaryotic member of the protein family of putative ATPases involved in membrane functions, cell cycle control and gene expression. Journal of Bacteriology **175**: 1344-1351.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from SDS and acid/urea polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proceedings of National Academy of Sciences, USA **76**: 4350-4354.

Tran, L., Wu, X.-C. and Wong, S.-L. (1991). Cloning and expression of a novel proteases gene encoding an extracellular neutral protease from *Bacillus subtilis*. Journal of Bacteriology **173**: 6364-6372.

Ullmann, A., Jacob, F. and Monod, J. (1967). Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of *Escherichia coli*. Journal of Molecular Biology **24**: 339.

Vallee, B. L., Stein, E. A., Sumerwell, W. N. and Fischer, E. H. (1959). Metal content of α-amylases of various origins. Journal of Biological Chemistry **234**: 2901-2905.

van Dijl, J. M., Bolhuis, A., Tjalsma, H., de Jong, A., Venema, G. and Bron, S. (1995). Characterization of signal peptidases in *Bacillus*. Proceedings of the 8th International Conference on Bacilli, Stanford University, USA..

van Dijl, J, M., de Jong, A., Vehmaanpera, J., Venema, G. and Bron, S. (1992). Signal peptidase I of *Bacillus subtilis* patterns of conserved amino acids in prokaryotic and eukaryotic type I signal peptidases. EMBO Journal **11**: 2819-2828.

van Dijl, J. M., de Jong, A., Venema, G. and Bron, S. (1995). Identification of the potential active site of the signal peptidase SipS of *Bacillus subtilis*. Journal of Biological Chemistry **270**: 3611-3618.

Violet, M. and Meunier, J.-C. (1989). Kinetic studies of the irreversible thermal inactivation of *Bacillus licheniformis* α-amylase. Biochemical Journal **263**: 665-670.

von Heijne, G. (1985). Signal sequences: The limits of variation. Journal of Molecular Biology. **184**: 99-105.

von Heijne, G. (1989). The structure of signal peptides from bacterial lipoproteins. Protein Engineering **2**: 531-534.

Vos, P., van Asseldonk, M., van Jeveren, F., Seizen, R., Simons, G. and de Vos, W. M. (1989). A maturation protein is required for production of active forms of *Lactococcus lactis* SK11 serine proteinase located in or secreted from the cell envelope. Journal of Bacteriology **171**: 2795-2802.

Wallerstein, M. (1909). US Patent No. 905029.

Walter, P. and Blobel, G. (1982). Signal recognotion particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature. **299**: 691-698.

Wandersman, C. (1990). Secretion, processing and activation of bacterial extracellular proteases. Molecular Microbiology **3**: 1825-1831.

Ward, J. B. (1973). The chain lengths of glycans in bacterial cell walls. Biochemical Journal **133**: 395-398.

Ward, J. B. (1981). Teichoic and teichuronic acids: Biosynthesis, assembly and location. Microbiological Reviews **45**: 211-243.

Watanabe, M. and Blobel, G. (1989). Site-specific antibodies against the PrlA (SecY) protein of *Escherichia coli* inhibit protein export by interfering with plasma membrane binding of preproteins. Proceedings of National Academy of Sciences, USA **86**: 1895-1899.

Weickert, J. and Chambliss, G. H. (1990). Site-directed mutagenesis of a catabolite repression operator in *Bacillus subtilis*. Proceedings of National Academy of Sciences, USA **87**: 6238-6242.

Wetzstein, M., Volker, U., Dedio, J., Lobau, S., Zuber, U., Schiesswohl, M., Herget, C., Hecker, M. and Schumann, W. (1992). Cloning, sequencing, and molecular analysis of the *dnaK* locus from *Bacillus subtilis*. Journal of Bacteriology 174: 3300-3310.

Wickner, W., Driessen, A. J. M. and Hartl, F. -U. (1991). The enzymology of protein translocation across the *Eschericia coli* plasma membrane. Annual Reviews in Biochemistry **60**: 101-124.

Wiedmann, M., Kurzchalia, T. V., Bielka, H. and Rapoport, T. A. (1987). Direct probing of the interaction between the signal sequence of the nascent preprolactin and the signal recognition particle by specific cross-linking. Journal of Cell Biology. **104**: 201-208.

Wild, J., Altman, E., Yura, T. and Gross, C. A. (1992). DnaK and DnaJ heat shock proteins participate in protein export in *Eschericia coli*. Genes and Development **6**: 1165-1172.

Wilhelm, M. and Hollenberg, C. P. (1984). Selective cloning of *Bacillus subtilis* xylose isomerase and xylulokinase genes in *Escherichia coli* by IS-5 mediated expression. EMBO Journal **3**: 2555-2560.

Wong, S.-L, Price, C. W., Goldfarb, D. S and Doi, R. H (1984). The subtilisin E gene of *Bacillus subtilis* is transcribed from a sigma-37 promoter *in vivo*. Proceedings of National Academy of Sciences, USA **81**: 1184-1188.

Wong, S.-L. and Doi, R. H. (1986). Determination of the signal peptide cleavage site in the preprosubtilisin of *Bacillus subtilis*. Journal of Biological Chemistry **261**: 10176-10181.

Wright, J. and Heckels, J. E. (1975). The teichuronic acid of *Bacillus subtilis* W23 grown in a chemostat under phosphate limitation. Biochemical Journal **147**: 186-189.

Wu, H. C. and Hayashi, S. (1986). Lipoprotein secretion in bacteria. In: Microbiology-1986 (Leive, L., ed). American Society for Microbiology, Washington DC, USA.

Wu, X., Lee, W., Tran, L. and Wong, S. (1991). Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. Journal of Bacteriology **173**: 4952-4958.

Wu, X. C., Nathoo, S., Pang, S. H., Carne, T. and Wong, S. L. (1990). Cloning, genetic organization, and characterization of a structural gene encoding bacillopeptidase F from *Bacillus subtilis*. Journal of Biological Chemistry **265**: 6845-6850.

Yamanaka, H., Kameyama, M., Baba, T., Fujii, Y. and Okamoto, K. (1994). Maturation pathway of *Escherichia coli* heat-stable enterotoxin I: requirement of DsbA for disulphide bond formation. Journal of Bacteriology **176**: 2906-2913.

Yang, M. Y., Ferrari, E. and Henner, D. J. (1984). Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create *in vitro*-derived deletion mutants. Journal of Bacteriology **160**: 15-21.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**: 103-119.

Yoneda, Y., Graham, S. and Young, F. E. (1979). Cloning of a foreign gene coding for α-amylase in *Bacillus subtilis*. Biomedical and Biophysical Research Communications **91**: 1556-1564.

Yoneyama, T., Araki, Y. and Ito, E. (1984). The primary structure of teichuronic acid in *Bacillus subtilis* AHU 1031. European Journal of Biochemistry **141**: 83-89.

Yuan, G. and Wong, S. L. (1995). Isolation and characterization of *Bacillus subtilis groE* regulatory mutants: evidence for *orf39* in the *dnaK* operon as a repressor gene in regulating the expression of both *groE* and *dnaK*. Journal of Bacteriology **177**: 6462-6468.

Zhu, X., Ohta, Y., Jordan, F. and Inouye, M. (1989). Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature **339**: 483-484.

Zuber, U. and Schumann, W. (1994). CIRCE, a novel heat shock element involved in regulation of heat shock *dnaK* of *Bacillus subtilis*. Journal of Bacteriology **176**: 1359-1363.

APPENDICES

APPENDIX 1 - GENE AND PROTEIN SEQUENCES

```
1     ATGAAACAAC AAAAACGGCT TACGCCCGA TTGCTGACGC TGTTATTTGC GCTCATCTTC
       M  K  Q  Q   K  R  L   Y  A  R   L  L  T  L   L  F  A   L  I  F

61    TTGCTGCCTC ATTCTGCAGC AGCGGCGGCA AATCTTAATG GGACGCTGAT GCAGTATTTT
       L  L  P  H   S  A  A   A  A  A   N  L  N  G   T  L  M   Q  Y  F

121   GAATGGTACA TGCCCAATGA CGGCCAACAT TGGAAGCGTT TGCAAAACGA CTCGGCATAT
       E  W  Y  M   P  N  D   G  Q  H   W  K  R  L   Q  N  D   S  A  Y

181   TTGGCTGAAC ACGGTATTAC TGCCGTCTGG ATTCCCCCGG CATATAAGGG AACGAGCCAA
       L  A  E  H   G  I  T   A  V  W   I  P  P  A   Y  K  G   T  S  Q

241   GCGGATGTGG GCTACGGTGC TTACGACCTT TATGATTTAG GGGAGTTTCA TCAAAAAGGG
       A  D  V  G   Y  G  A   Y  D  L   Y  D  L  G   E  F  H   Q  K  G

301   ACGGTTCGGA CAAAGTACGG CACAAAAGGA GAGCTGCAAT CTGCGATCAA AAGTCTTCAT
       T  V  R  T   K  Y  G   T  K  G   E  L  Q  S   A  I  K   S  L  H

361   TCCCGCGACA TTAACGTTTA CGGGGATGTG GTCATCAACC ACAAAGGCGG CGCTGATGCG
       S  R  D  I   N  V  Y   G  D  V   V  I  N  H   K  G  G   A  D  A

421   ACCGAAGATG TAACCGCGGT TGAAGTCGAT CCCGCTGACC GCAACCGCGT AATTTCAGGA
       T  E  D  V   T  A  V   E  V  D   P  A  D  R   N  R  V   I  S  G

481   GAACACCTAA TTAAAGCCTG GACACATTTT CATTTTCCGG GGCGCGGCAG CACATACAGC
       E  H  L  I   K  A  W   T  H  F   H  F  P  G   R  G  S   T  Y  S

541   GATTTTAAAT GGCATTGGTA CCATTTTGAC GGAACCGATT GGGACGAGTC CCGAAAGCTG
       D  F  K  W   H  W  Y   H  F  D   G  T  D  W   D  E  S   R  K  L

601   AACCGCATCT ATAAGTTTCA AGGAAAGGCT TGGGATTGGG AAGTTTCCAA TGAAAACGGC
       N  R  I  Y   K  F  Q   G  K  A   W  D  W  E   V  S  N   E  N  G

661   AACTATGATT ATTTGATGTA TGCCGACATC GATTATGACC ATCCTGATGT CGCAGCAGAA
       N  Y  D  Y   L  M  Y   A  D  I   D  Y  D  H   P  D  V   A  A  E

721   ATTAAGAGAT GGGGCACTTG GTATGCCAAT GAACTGCAAT TGGACGGTTT CCGTCTTGAT
       I  K  R  W   G  T  W   Y  A  N   E  L  Q  L   D  G  F   R  L  D

781   GCTGTCAAAC ACATTAAATT TTCTTTTTTG CGGGATTGGG TTAATCATGT CAGGGAAAAA
       A  V  K  H   I  K  F   S  F  L   R  D  W  V   N  H  V   R  E  K

841   ACGGGGAAGG AAATGTTTAC GGTAGCTGAA TATTGGCAGA ATGACTTGGG CGCGCTGGAA
       T  G  K  E   M  F  T   V  A  E   Y  W  Q  N   D  L  G   A  L  E

901   AACTATTTGA ACAAAACAAA TTTTAATCAT TCAGTGTTTG ACGTGCCGCT TCATTATCAG
       N  Y  L  N   K  T  N   F  N  H   S  V  F  D   V  P  L   H  Y  Q
```

```
961   TTCCATGCTG CATCGACACA GGGAGGCGGC TATGATATGA GGAAATTGCT GAACGGTACG
        F  H  A  A   S  T  Q   G  G  G   Y  D  M  R   K  L  L   N  G  T

1021  GTCGTTTCCA AGCATCCGTT GAAATCGGTT ACATTTGTCG ATAACCATGA TACACAGCCG
        V  V  S  K   H  P  L   K  S  V   T  F  V  D   N  H  D   T  Q  P

1081  GGGCAATCGC TTGAGTCGAC TGTCCAAACA TGGTTTAAGC CGCTTGCTTA CGCTTTTATT
        G  Q  S  L   E  S  T   V  Q  T   W  F  K  P   L  A  Y   A  F  I

1141  CTCACAAGGG AATCTGGATA CCCTCAGGTT TTCTACGGGG ATATGTACGG GACGAAAGGA
        L  T  R  E   S  G  Y   P  Q  V   F  Y  G  D   M  Y  G   T  K  G

1201  GACTCCCAGC GCGAAATTCC TGCCTTGAAA CACAAAATTG AACCGATCTT AAAAGCGAGA
        D  S  Q  R   E  I  P   A  L  K   H  K  I  E   P  I  L   K  A  R

1261  AAACAGTATG CGTACGGAGC ACAGCATGAT TATTTCGACC ACCATGACAT TGTCGGCTGG
        K  Q  Y  A   Y  G  A   Q  H  D   Y  F  D  H   H  D  I   V  G  W

1321  ACAAGGGAAG GCGACAGCTC GGTTGCAAAT TCAGGTTTGG CGGCATTAAT AACAGACGGA
        T  R  E  G   D  S  S   V  A  N   S  G  L  A   A  L  I   T  D  G

1381  CCCGGTGGGG CAAAGCGAAT GTATGTCGGC CGGCAAAACG CCGGTGAGAC ATGGCATGAC
        P  G  G  A   K  R  M   Y  V  G   R  Q  N  A   G  E  T   W  H  D

1441  ATTACCGGAA ACCGTTCGGA GCCGGTTGTC ATCAATTCGG AAGGCTGGGG AGAGTTTCAC
        I  T  G  N   R  S  E   P  V  V   I  N  S  E   G  W  G   E  F  H

1501  GTAAACGGCG GGTCGGTTTC AATTTATGTT CAAAGATAGA AGAGCAGAGA GGACGGATTT
        V  N  G  G   S  V  S   I  Y  V   Q  R  .

1561  CCTGAAGGAA ATCCGTTTTT TTATTTTGCC CGTCTTATAA ATTTCTTTGA TTACATTTTA

1621  TAATTAATTT TAACAAAGTG TCATCAGCCC TCAGGAAGGA CTTGCTGACA GTTTGAATCG

1681  CATAGGTAAG GCGGGGATGA AATGGCAACG TTATCTGATG TAGCAAAGAA AGCAAATGTG

1741  TCGAAAATGA CGGTATCGCG GGTGATCAAT CATCCTGAGA CTGTGACGGA TGAATTGAAA

1801  AAGCTT
```

The nucleotide sequence of the *amyL* gene and encoded amino acids (Gray *et al.*, 1986).

```
1     ATGATTCAAA AACGAAAGCG GACAGTTTCG TTCAGACTTG TGCTTATGTG CACGCTGTTA
        M  I  Q  K   R  K  R   T  V  S   F  R  L  V   L  M  C   T  L  L

61    TTTGTCAGTT TGCCGATTAC AAAAACATCA GCCGTAAATG GCACGCTGAT GCAGTATTTT
        F  V  S  L   P  I  T   K  T  S   A  V  N  G   T  L  M   Q  Y  F

121   GAATGGTATA CGCCGAACGA CGGCCAGCAT TGGAAACGAT TGCAGAATGA TGCGGAACAT
        E  W  Y  T   P  N  D   G  Q  H   W  K  R  L   Q  N  D   A  E  H

181   TTATCGGATA TCGGAATCAC TGCCGTCTGG ATTCCTCCCG CATACAAAGG ATTGAGCCAA
        L  S  D  I   G  I  T   A  V  W   I  P  P  A   Y  K  G   L  S  Q

241   TCCGATAACG GATACGGACC TTATGATTTG TATGATTTAG GAGAATTCCA GCAAAAAGGG
        S  D  N  G   Y  G  P   Y  D  L   Y  D  L  G   E  F  Q   Q  K  G

301   ACGGTCAGAA CGAAATACGG CACAAAATCA GAGCTTCAAG ATGCGATCGG CTCACTGCAT
        T  V  R  T   K  Y  G   T  K  S   E  L  Q  D   A  I  G   S  L  H

361   TCCCGGAACG TCCAAGTATA CGGAGATGTG GTTTTGAATC ATAAGGCTGG TGCTGATGCA
        S  R  N  V   Q  V  Y   G  D  V   V  L  N  H   K  A  G   A  D  A

421   ACAGAAGATG TAACTGCCGT CGAAGTCAAT CCGGCCAATA GAAATCAGGA AACTTCGGAG
        T  E  D  V   T  A  V   E  V  N   P  A  N  R   N  Q  E   T  S  E

481   GAATATCAAA TCAAAGCGTG GACGGATTTT CGTTTTCCGG GCCGTGGAAA CACGTACAGT
        E  Y  Q  I   K  A  W   T  D  F   R  F  P  G   R  G  N   T  Y  S

541   GATTTTAAAT GGCATTGGTA TCATTTCGAC GGAGCGGACT GGGATGAATC CCGGAAGATC
        D  F  K  W   H  W  Y   H  F  D   G  A  D  W   D  E  S   R  K  I

601   AGCCGCATCT TTAAGTTTCG TGGGGAAGGA AAAGCGTGGG ATTGGGAAGT ATCAAGTGAA
        S  R  I  F   K  F  R   G  E  G   K  A  W  D   W  E  V   S  S  E

661   AACGGCAACT ATGACTATTT AATGTATGCT GATGTTGACT ACGACCACCC TGATGTCGTG
        N  G  N  Y   D  Y  L   M  Y  A   D  V  D  Y   D  H  P   D  V  V

721   GCAGAGACAA AAAAATGGGG TATCTGGTAT GCGAATGAAC TGTCATTAGA CGGCTTCCGT
        A  E  T  K   K  W  G   I  W  Y   A  N  E  L   S  L  D   G  F  R

781   ATTGATGCCG CCAAACATAT TAAATTTTCA TTTCTGCGTG ATTGGGTTCA GGCGGTCAGA
        I  D  A  A   K  H  I   K  F  S   F  L  R  D   W  V  Q   A  V  R

841   CAGGCGACGG GAAAAGAAAT GTTTACGGTT GCGGAGTATT GGCAGAATAA TGCCGGGAAA
        Q  A  T  G   K  E  M   F  T  V   A  E  Y  W   Q  N  N   A  G  K

901   CTCGAAAACT ACTTGAATAA AACAAGCTTT AATCAATCCG TGTTTGATGT TCCGCTTCAT
        L  E  N  Y   L  N  K   T  S  F   N  Q  S  V   F  D  V   P  L  H
```

961 TTCAATTTAC AGGCGGCTTC CTCACAAGGA GGCGGATATG ATATGAGGCG TTTGCTGGAC
F N L Q A A S S Q G G G Y D M R R L L D

1021 GGTACCGTTG TGTCCAGGCA TCCGGAAAAG GCGGTTACAT TTGTTGAAAA TCATGACACA
G T V V S R H P E K A V T F V E N H D T

1081 CAGCCGGGAC AGTCATTGGA ATCGACAGTC CAAACTTGGT TTAAACCGCT TGCATACGCC
Q P G Q S L E S T V Q T W F K P L A Y A

1141 TTTATTTTGA CAAGAGAATC CGGTTATCCT CAGGTGTTCT ATGGGGATAT GTACGGGACA
F I L T R E S G Y P Q V F Y G D M Y G T

1201 AAAGGGACAT CGCCAAAGGA AATTCCCTCA CTGAAAGATA ATATAGAGCC GATTTTAAAA
K G T S P K E I P S L K D N I E P I L K

1261 GCGCGTAAGG AGTACGCATA CGGGCCCCAG CACGATTATA TTGACCACCC GGATGTGATC
A R K E Y A Y G P Q H D Y I D H P D V I

1321 GGATGGACGA GGGAAGGTGA CAGCTCCGCC GCCAAATCAG GTTTGGCCGC TTTAATCACG
G W T R E G D S S A A K S G L A A L I T

1381 GACGGACCCG GCGGATCAAA GCGGATGTAT GCCGGCCTGA AAAATGCCGG CGAGACATGG
D G P G G S K R M Y A G L K N A G E T W

1441 TATGACATAA CGGGCAACCG TTCAGATACT GTAAAAATCG GATCTGACGG CTGGGGAGAG
Y D I T G N R S D T V K I G S D G W G E

1501 TTTCATGTAA ACGATGGGTC CGTCTCCATT TATGTTCAGA AATAAGGTAA TAAAAAAACA
F H V N D G S V S I Y V Q K .

1561 CCTCCAAGCT GAGTGCGGGT ATCAGCTTGG AGGTGCGTTT ATTTTTTCAG CCGTATGACA

1621 AGGTCGGCAT CAGGTGTGAC AAATACGGTA TGCTGGCTGT CATAGGTGAC AAATCCGGGT

1681 TTTGCGCCGT TTGGCTTTTT CACATGTCTG ATTTTTGTAT AATCAACAGG CACGGAGCCG

1741 GAATCTTTCG CCTTGGAAAA ATAAGCGGCG ATCGTAGCTG CTTCCAATAT GGATTGTTCA

1801 TCGGGATCGC TGCTTTTAAT CACAACGTGG GATC

The nucleotide sequence of the *amyQ* gene and encoded amino acids (Takkinen *et al.*, 1983).

```
1     GTGCTAACGT TTCACCGCAT CATTCGAAAA GGGTGGGTGT TCCTGCTCGC GTTTTGGCTC
       M  L  T  F  H  R  I  I  R  K  G  W  V  F  L  L  A  F  W  L

61    ACTGCCTCGC TGTTCTGCCC GACAGGACAG CCCGCCAAGG CTGCCGCACC GTTTAACGGC
       T  A  S  L  F  C  P  T  G  Q  P  A  K  A  A  A  P  F  N  G

121   ACCATGATGC AGTATTTTGA ATGGTACTTG CCGGATGATG GCACGTTATG GACCAAAGTG
       T  M  M  Q  Y  F  E  W  Y  L  P  D  D  G  T  L  W  T  K  V

181   GCCAATGAAG CCAACAACTT ATCCAGCCTT GGCATCACCG CTCTTTGGCT GCCGCCCGCT
       A  N  E  A  N  N  L  S  S  L  G  I  T  A  L  W  L  P  P  A

241   TATAAAGGAA CAAGCCGCAG CGACGTAGGG TACGGAGTAT ACGACTTGTA TGACCTCGGC
       Y  K  G  T  S  R  S  D  V  G  Y  G  V  Y  D  L  Y  D  L  G

301   GAATTCAATC AAAAAGGGAC CGTCCGCACA AAATACGGAA CAAAAGCTCA ATATCTTCAA
       E  F  N  Q  K  G  T  V  R  T  K  Y  G  T  K  A  Q  Y  L  Q

361   GCCATTCAAG CCGCCCACGC CGCTGGAATG CAAGTGTACG CCGATGTCGT GTTCGACCAT
       A  I  Q  A  A  H  A  A  G  M  Q  V  Y  A  D  V  V  F  D  H

421   AAAGGCGGCG CCGACGGCAC GGAATGGGTG GACGCCGTCG AAGTCAATCC GTCCGACCGC
       K  G  G  A  D  G  T  E  W  V  D  A  V  E  V  N  P  S  D  R

481   AACCAAGAAA TCTCGGGCAC CTATCAAATC CAAGCATGGA CGAAATTTGA TTTTAACGGG
       N  Q  E  I  S  G  T  Y  Q  I  Q  A  W  T  K  F  D  F  N  G

541   CGGGGCAACA CCTACTCCAG CTTTAAGTGG CGCTGGTACC ATTTTGACGG CGTTGACTGG
       R  G  N  T  Y  S  S  F  K  W  R  W  Y  H  F  D  G  V  D  W

601   GACGAAAGCC GAAAATTAAG CCGCATTTAC AAATTCCGCG GCATCGGCAA AGCGTGGGAT
       D  E  S  R  K  L  S  R  I  Y  K  F  R  G  I  G  K  A  W  D

661   TGGGAAGTAG ACACGGAAAA CGGAAACTAT GACTACTTAA TGTATGCCGA CCTTGATATG
       W  E  V  D  T  E  N  G  N  Y  D  Y  L  M  Y  A  D  L  D  M

721   GATCATCCCG AAGTCGTGAC CGAGCTGAAA AACTGGGGGA AATGGTATGT CAACACAACG
       D  H  P  E  V  V  T  E  L  K  N  W  G  K  W  Y  V  N  T  T

781   AACATTGATG GGTTCCGGCT TGATGCTGTC AAGCATATTA AGTTCAGTTT TTTCCTGAT
       N  I  D  G  F  R  L  D  A  V  K  H  I  K  F  S  F  F  P  D

841   TGGTTGTCGT ATGTGCGTTC TCAGACTGGC AAGCCGCTAT TTACCGTCGG GGAATATTGG
       W  L  S  Y  V  R  S  Q  T  G  K  P  L  F  T  V  G  E  Y  W

901   AGCTATGACA TCAACAAGTT GCACAATTAC ATTACGAAAA CAAGCGGAAC GATGTCTTTG
       S  Y  D  I  N  K  L  H  N  Y  I  T  K  T  S  G  T  M  S  L
```

961 TTTGATGCCC CGTTACACAA CAAATTTTAT ACCGCTTCCA AATCAGGGGG CGCATTTGAT
F D A P L H N K F Y T A S K S G G A F D

1021 ATGAGCACGT TAATGAACAA TACTCTCATG AAAGATCAAC CGACATTGGC CGTCACCTTC
M S T L M N N T L M K D Q P T L A V T F

1081 GTTGATAATC ATGACACCGA ACCCGGCCAA GCGCTGCAGT CATGGGTCGA CCCATGGTTC
V D N H D T E P G Q A L Q S W V D P W F

1141 AAACCGTTGG CTTACGCCTT TATTCTAACT CGGCAGGAAG GATACCCGTG CGTCTTTTAT
K P L A Y A F I L T R Q E G Y P C V F Y

1201 GGTGACTATT ATGGCATTCC ACAATATAAC ATTCCTTCAC TGAAAAGCAA AATCGATCCG
G D Y Y G I P Q Y N I P S L K S K I D P

1261 CTCCTCATCG CGCGCAGGGA TTATGCTTAT GGAACGCAAC ATGATTATCT TGATCACTCC
L L I A R R D Y A Y G T Q H D Y L D H S

1321 GACATCATCG GGTGGACAAG GGAAGGCGTT ACCGAAAAAC CAGGATCCGG ACTGGCCGCA
D I I G W T R E G V T E K P G S G L A A

1381 CTGATCACCG ATGGGCCGGG AGGAAGCAAA TGGATGTACG TTGGCAAACA ACACGCCGGA
L I T D G P G G S K W M Y V G K Q H A G

1441 AAAGTGTTCT ATGACCTTAC CGGCAACCGG AGTGACACCG TCACCATCAA CAGTGATGGA
K V F Y D L T G N R S D T V T I N S D G

1501 TGGGGGGAAT TCAAAGTCAA TGGCGGTTCG GTTTCGGTTT GGGTTCCTAG AAAAACGACC
W G E F K V N G G S V S V W V P R K T T

1561 GTTTCTACCA TCGCTCGGCC GATCACAACC CGACCGTGGA CTGGTGAATT CGTCCGTTGG
V S T I A R P I T T R P W T G E F V R W

1621 ACCGAACCAC GGTTGGTGGC ATGGCCTTGA TGCCTGCGAT CGCGTTGTAA AGACATTCCG
T E P R L V A W P .

1681 CTCTATCATT GAGGCAAAAA ACATGGCCTT GTCCGCCATG AATGGCGGCA CAAGGCCGTG

1741 TTTGATGTTA CCATCCATTT GCTTGCTTCA ACTTTTCCTT CGACGGCGTT TCGTAGCGGA

1801 TGTGCGTGTC GATGTCGGTC ACGTAATACC CGCCGCCGAC CGCGTATTGC CCGCGAAGCG

1861 CGCGTCATAC

The nucleotide sequence of the *amyS* gene and encoded amino acids (Gray *et al.*, 1986).

```
1     ATGAAACAAC AAAAACGGCT TTACGCCCGA TTGCTGACGC TGTTATTTGC GCTCATCTTC
        M  K  Q  Q   K  R  L   Y  A  R   L  L  T  L   L  F  A   L  I  F

61    TTGCTGCCTC ATTCTGCAGC AGCGGCGGCA AATCTTAATG GGACGCTGAT GCAGTATTTT
        L  L  P  H   S  A  A   A  A  A   N  L  N  G   T  L  M   Q  Y  F

121   GAATGGTACA TGCCCAATGA CGGCCAACAT TGGAAGCGTT TGCAAAACGA CTCGGCATAT
        E  W  Y  M   P  N  D   G  Q  H   W  K  R  L   Q  N  D   S  A  Y

181   TTGGCTGAAC ACGGTATTAC TGCCGTCTGG ATTCCCCCGG CATATAAGGG AACAAGCCGC
        L  A  E  H   G  I  T   A  V  W   I  P  P  A   Y  K  G   T  S  R

241   AGCGACGTAG GGTACGGTGC TTACGACCTT TATGATTTAG GGGAGTTTCA TCAAAAAGGG
        S  D  V  G   Y  G  A   Y  D  L   Y  D  L  G   E  F  H   Q  K  G

301   ACGGTTCGGA CAAAGTACGG CACAAAAGGA GAGCTGCAAT CTGCGATCAA AAGTCTTCAT
        T  V  R  T   K  Y  G   T  K  G   E  L  Q  S   A  I  K   S  L  H

361   TCCCGCGACA TTAACGTTTA CGGGGATGTG GTCATCAACC ACAAAGGCGG CGCTGATGCG
        S  R  D  I   N  V  Y   G  D  V   V  I  N  H   K  G  G   A  D  A

421   ACCGAAGATG TAACCGCGGT TGAAGTCGAT CCCGCTGACC GCAACCGCGT AATTTCAGGA
        T  E  D  V   T  A  V   E  V  D   P  A  D  R   N  R  V   I  S  G

481   GAACACCTAA TTAAAGCCTG GACACATTTT CATTTTCCCG GGCGGGGCAA CACCTACTCC
        E  H  L  I   K  A  W   T  H  F   H  F  P  G   R  G  N   T  Y  S

541   AGCTTTAAGT GGCGCTGGTA CCATTTTGAC GGAACCGATT GGGACGAGTC CCGAAAGCTG
        S  F  K  W   R  W  Y   H  F  D   G  T  D  W   D  E  S   R  K  L

601   AACCGCATCT ATAAGTTTCA AGGAAAGGCT TGGGATTGGG AAGTTTCCAA TGAAAACGGC
        N  R  I  Y   K  F  Q   G  K  A   W  D  W  E   V  S  N   E  N  G

661   AACTATGATT ATTTGATGTA TGCCGACATC GATTATGACC ATCCTGATGT CGCAGCAGAA
        N  Y  D  Y   L  M  Y   A  D  I   D  Y  D  H   P  D  V   A  A  E

721   ATTAAGAGAT GGGGCACTTG GTATGCCAAT GAACTGCAAT TGGACGGTTT CCGTCTTGAT
        I  K  R  W   G  T  W   Y  A  N   E  L  Q  L   D  G  F   R  L  D

781   GCTGTCAAAC ACATTAAATT TTCTTTTTTG CGGGATTGGG TTAATCATGT CAGGGAAAAA
        A  V  K  H   I  K  F   S  F  L   R  D  W  V   N  H  V   R  E  K

841   ACTGGCAAGC CGCTATTTAC CGTCGGGGAA TATTGGAGCT ATGACATCAA CAAGTTGCAC
        T  G  K  P   L  F  T   V  G  E   Y  W  S  Y   D  I  N   K  L  H

901   AATTACATTA CGAAAACAAG CGGAACGATG TCTTTGTTTG ATGCCCCGTT ACACAACAAA
        N  Y  I  T   K  T  S   G  T  M   S  L  F  D   A  P  L   H  N  K
```

```
961  TTTTATACCG CTTCCAAATC AGGAGGCGGC TATGATATGA GGAAATTGCT GAACGGTACG
       F  Y  T  A   S  K  S   G  G  G   Y  D  M  R   K  L  L   N  G  T

1021 GTCGTTTCCA AGCATCCGTT GAAATCGGTT ACATTTGTCG ATAACCATGA TACACAGCCG
       V  V  S  K   H  P  L   K  S  V   T  F  V  D   N  H  D   T  Q  P

1081 GGGCAATCGC TTGAGTCGAC TGTCCAAACA TGGTTTAAGC CGCTTGCTTA CGCTTTTATT
       G  Q  S  L   E  S  T   V  Q  T   W  F  K  P   L  A  Y   A  F  I

1141 CTCACAAGGG AATCTGGATA CCCTCAGGTT TTCTACGGGG ATATGTACGG GACGAAAGGG
       L  T  R  E   S  G  Y   P  Q  V   F  Y  G  D   M  Y  G   T  K  G

1201 ACATCGCCAA AGGAAATTCC CTCACTGAAA CACAAAATTG AACCGATCTT AAAAGCGAGA
       T  S  P  K   E  I  P   S  L  K   H  K  I  E   P  I  L   K  A  R

1261 AAACAGTATG CGTACGGAGC ACAGCATGAT TATTTCGACC ACCATGACAT TGTCGGCTGG
       K  Q  Y  A   Y  G  A   Q  H  D   Y  F  D  H   H  D  I   V  G  W

1321 ACAAGGGAAG GCGACAGCTC GGTTGCAAAT TCAGGTTTGG CGGCATTAAT AACAGACGGA
       T  R  E  G   D  S  S   V  A  N   S  G  L  A   A  L  I   T  D  G

1381 CCCGGTGGGG CAAAGCGAAT GTACGTTGGC AAACAACACG CCGGAAAAGT GTTCTATGAC
       P  G  G  A   K  R  M   Y  V  G   K  Q  H  A   G  K  V   F  Y  D

1441 CTTACCGGCA ACCGGAGTGA CACCGTCACC ATCAACAGTG ATGGATGGGG GGAATTCAAA
       L  T  G  N   R  S  D   T  V  T   I  N  S  D   G  W  G   E  F  K

1501 GTCAATGGCG GTTCGGTTTC GGTTTGGGTT CCTAGAAAAA CGACCGTTTC TACCATCGCT
       V  N  G  G   S  V  S   V  W  V   P  R  K  T   T  V  S   T  I  A

1561 CGGCCGATCA CAACCCGACC GTGGACTGGT GAATTCGTCC GTTGGACCGA ACCACGGTTG
       R  P  I  T   T  R  P   W  T  G   E  F  V  R   W  T  E   P  R  L

1621 GTGGCATGGC CTTAGAAGAG CAGAGAGGAC GGATTTCCTG AAGGAAATCC GTTTTTTTAT
       V  A  W  P   .

1681 TTTGCCCGTC TTATAAATTT CTTTGATTAC ATTTTATAAT TAATTTTAAC AAAGTGTCAT

1741 CAGCCCTCAG GAAGGACTTG CTGACAGTTT GAATCGCATA GGTAAGGCGG GGATGAAATG

1801 GCAACGTTAT CTGATGTAGC AAAGAAAGCA AATGTGTCGA AAATGACGGT ATCGCGGGTG

1861 ATCAATCATC CTGAGACTGT GACGGATGAA TTGAAAAAGC TT
```

The nucleotide sequence of the *amyLQS50* gene and encoded amino acids.

APPENDIX 2 - OLIGONUCLEOTIDE PRIMER SEQUENCES

Primer	Sequence (5' to 3')
KS-3	GGGGAGATCTGCTCTTCCGCTTCCTCG
KS-4	GGGGAGATCTAATATTGAAAAAGGAAGAG
LQS50/1	GGGGGAATTCTTGCTGCCTCATTCTGCAGCAGCG
LQS50/2	GTACCCTACGTCGCTGCGGCTTGTTCCCTTATATGCCGGGGG
LQS50/3	ACAAGCCGCAGCGACGTAGGGTACGGTGCTTACGACCTTTATGATT TAG
LQS50/4	GGGGGGAGCTCTTCAACCGCGGTTACATCTTCGG
LQS50/5	GGGGGGAGCTCTGTAACCGCGGTTGAAGTCGATC
LQS50/6	CCAGCGCCACTTAAAGCTGGAGTAGGTGTTGCCCCGCCCGGGAAAA TGAAAATGTGTCCAGG
LQS50/7	CCCGGGCGGGGCAACACCTACTCCAGCTTTAAGTGGCGCTGGTACC ATTTTGACGGAACCGATTG
LQS50/8	GGGGGGGATCCTGGTCATAATCGATGTCGGCATAC
LQS50/9	GGGGGGGATCCATGCCGACATCGATTATGACCATC
LQS50/10	GTAAATAGCGGCTTGCCAGTTTTTTCCCTGACATGATTAAG
LQS50/11	TTAATCATGTCAGGGAAAAAACTGGCAAGCCGCTATTTAC
LQS50/12	CTCATATCATAGCCGCCTCCTGATTTGGAAGCGGTATAAAATTTG
LQS50/13	TTTATACCGCTTCCAAATCAGGAGGCGGCTATGATATGAG
LQS50/14	CCATGTTTGGACAGTCGACTCAAGC
LQS50/15	GGGGCAATCGCTTGAGTCGACTGTCCAAAC
LQS50/16	TTCCTTTGGCGATGTCCCTTTCGTCCCGTACATATCCC

LQS50/17	GGGACATCGCCAAAGGAAATTCCCTCACTGAAACACAAAATTG
LQS50/18	GGGGGGCATGCCTGTGCTCCGTACGCGTACTGTTTTCTG
LQS50/19	GGGGGGGATCCCGAGAAAACAGTACGCGTACGGAGCAC
LQS50/20	TGTTGTTTGCCAACGTACATTCGCTTTGCCCCACCGGGTC
LQS50/21	GACCCGGTGGGGCAAAGCGAATGTACGTTGGCAAACAACAC
LQS50/22	CCGTCCTCTCTGCTCTTCTAAGGCCATGCCACCAACCGTG
LQS50/23	CACGGTTGGTGGCATGGCCTTAGAAGAGCAGAGAGGACGG
LQS50/24	GACAGCAAGCTTTTTCAATTCATCCG
LQS50/25	CAGACGGACCCGGTGGGGCA
LQS50/26	ACGAAATTTATAAGACGGGC
M13 Forward	GTTTTCCCAGTCACGAC
M13 Reverse	CAGGAAACAGCTATGAC

APPENDIX 3 - MEDIA, BUFFERS AND REAGENTS

17 Amino acid solution (x100)
2 mg/ml solution of all naturally ocurring amino acids except methionine, cysteine and tyrosine.

BCG solution

x10 MM	100 ml
Trisodium citrate.$7H_2O$ (10% w/v)	10 ml
$MgSO_4$ (1 M)	2 ml
Salt mix	1 ml
Glucose (20% w/v)	20 ml
Distilled water	866 ml

Blocking buffer

Tris.Cl, pH 7.5	10 mM
EDTA	2 mM
NaCl	150 mM
BSA	0.5%
Triton X-100	0.05%
NaN_3	0.01%

Blotting buffer

Glycine	43.2 g
Tris.Cl, pH 7.5 (1 M)	75 ml
Methanol	600 ml
Distilled water	2375 ml

BTF solution

BCG solution	800 µl
EGTA (10mM)	100 µl
$MgCl_2$ (400mM)	100 µl

Coomassie blue solution

Kentacid blue	1.25 g
Methanol (50% v/v)	454 ml
Glacial acetic acid	46 ml

Denaturing solution	
NaCl	1.5 M
NaOH	0.5 M
Destain solution	
Methanol	30% v/v
Glacial acetic acid	10% v/v
DNA loading buffer (x10)	
Glycerol	8.2 M
EDTA	0.2 M
Bromophenol blue	72 mM
Hybridization solution	
x20 SSC	7.5 ml
x100 Denhardt's solution	250 µl
Denatured probe solution	45 µl
Distilled water	17.205 ml
KM stock solution	
x10 MM	10 ml
Trisodium citrate.$7H_2O$ (10% w/v)	1 ml
$MgSO_4$ (1 M)	200 µl
Distilled water	90 ml
KM-1 solution	
KM stock solution	48 ml
Glucose (20% w/v)	1 ml
Casamino acids (20% w/v)	50 µl
Yeast extract (20% w/v)	250 µl
$MgCl_2$ (1 mM)	1.5 µl

KM-2 solution

KM stock solution	48 ml
Glucose (20% w/v)	1 ml
Casamino acids (20% w/v)	50 µl
Yeast extract (20% w/v)	250 µl
Salt solution	100 µl
$CaCl_2$ (0.5 M)	100 µl
$MgCl_2$ (1M)	80 µl

L-broth

Tryptone	1.0% w/v
Yeast extract	0.5% w/v
NaCl	0.5% w/v

x2 Ligation buffer

Ligation buffer (x10, provided with ligase)	20 µl
Dithiothreitol (100 mM)	40 µl
ATP (10 mM)	20 µl
Distilled water	20 µl

Store at -20°C.

Lysis buffer

Tris.Cl, pH 7.2	10 mM
$MgCl_2$	25 mM
NaCl	200 mM
Lysozyme	5 mg/ml

MM solution (x10)

$(NH_4)_2SO_4$	2% w/v
KH_2PO_4	6% w/v
K_2HPO_4	14% w/v

Neutralizing solution

NaCl	1.5 M
Tris.Cl, pH 7.2	0.5 M

P1 buffer

Tris.Cl	50 mM
EDTA	10 mM
RNase A	100 µg/ml

Adjusted to pH 8.0.

P2 buffer

NaOH	200 mM
SDS	1% w/v

P3 buffer

Potassium acetate pH 5.5	3 M

Penassay agar

Antibiotic medium number 3 (Difco)	1.75% w/v
Agar	1.5% w/v

Penassay broth

Antibiotic medium number 3 (Difco)	1.75% w/v

PPO (Diphenyloxazole) solution

PPO	20% w/v

In glacial acetic acid.

Pre-hybridization solution

x20 SSC	7.5 ml
x100 Denhardt's solution	250 µl
Salmon testes DNA (1mg /ml)	500 µl
Distilled water	16.75 ml

QBT buffer

NaCl	750 mM
3-[N-morpholino]propanesulphonic acid (MOPS)	50 mM
Ethanol	15% v/v
Triton X-100	0.15%

pH adjusted to 7.0.

QC buffer	
NaCl	1 M
MOPS	50 mM
Ethanol	15% v/v
pH adjusted to 7.0.	
QF buffer	
NaCl	1.25 M
Tris.Cl	50 mM
Ethanol	15% v/v
pH adjusted to 8.5.	
Salt solution	
$CaCl_2$ (0.5 M)	400 µl
$FeCl_3$ (0.1 M)	200 µl
$MnCl_2$ (0.1 M)	200 µl
Distilled water	19.2 ml
x2 Sample buffer	
Tris.Cl, pH 6.8	0.125 M
Glycerol	20.0% v/v
SDS	4.0% w/v
2-Mercaptoethanol	10.0% v/v
Bromophenol blue	0.002 %
SDS-PAGE electrophoresis buffer	
Tris.Cl	0.394%
Glycine	1.44% w/v
SDS	0.1%
SOC	
Tryptone	2.0% w/v
Yeast extract	0.5% w/v
NaCl	10 mM
KCl	2.5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	20 mM

Sodium phosphate buffer

Phosphate buffer at the required pH was made by mixing x ml of solution A (0.2 M monobasic sodium phosphate) with y ml of solution B (0.2 M dibasic sodium phosphate), then making the volume up to 200ml with distilled water.

pH	x	y
5.7	93.5	5.7
6.1	85.0	15.0
6.7	56.5	43.5
7.0	39.0	61.0
7.2	28.0	72.0
7.3	23.0	77.0
7.7	10.5	89.5
8.0	5.3	94.7

Spizizen's minimal medium (SMM)

SMS (x10)	10 ml
Distilled water	67 ml
Ribose (10% w/v)	10 ml
Xylose (20% w/v)	5 ml
$MgSO_4.7H_20$ (2% w/v)	1 ml
17 Amino acid solution (x100)	1 ml
Trace element solution (x100)	1 ml
L-tryptophan (2 mg/ml)	5 ml

Spizizen's Minimal Salts (SMS)

$(NH_4)_2\ SO_4$	0.2% w/v
K_2HPO_4	1.4% w/v
KH_2PO_4	0.6% w/v
Na_3 citrate.$7H_20$	0.1% w/v

Can be made as a x10 stock.

SSC (x20)

NaCl	17.53% w/v
Trisodium citrate.$7H_20$	8.82% w/v

x10 STD

NaCl	9% w/v
Triton X-100	10% w/v
Sodium deoxycholate	5% w/v

x1 STD

x10 STD	10 ml
Tris.Cl, pH 7.2 (0.1 M)	10 ml
Distilled water	80 ml

STET buffer

Sucrose	8 g
Triton X-100	5 ml
EDTA, pH 8.0 (0.5 M)	10 ml
Tris.Cl, pH 8.0 (1 M)	5 ml

Make up to 100ml with distilled water.

TBE buffer

Tris.Cl	10.8% w/v
Boric acid	5.5% w/v
EDTA	0.93% w/v

TES

Tris.Cl, pH 8.0	25 mM
EDTA, pH 8.0	25 mM
Sucrose	0.3 M

TFBI

Potassium acetate	30 mM
$CaCl_2$	10 mM
KCl	100 mM
$MnCl_2$	50 mM
Glycerol	15% v/v

TFBII

$CaCl_2$	75 mM
KCl	10 mM
Glycerol	15% v/v
MOPS, pH 7.0	10 mM

Trace element solution (x100)

$MgCl_2.6H_2O$	1.25% w/v
$CaCl_2$	0.055% w/v
$FeCl_2.6H_2O$	0.135% w/v
$MnCl_2.4H_2O$	0.01% w/v
$ZnCl_2$	0.017% w/v
$CuCl_2.2H_2O$	0.0043% w/v
$CoCl_2.6H_2O$	0.006% w/v
$Na_2MoO_4.2H_2O$	0.006% w/v

Tris.Cl

Tizma base adjusted to the required pH with HCl.

Wash solution I

Tris.Cl, pH 7.5	10 mM
EDTA	2 mM
NaCl	150 mM
BSA	0.1%
Triton X-100	0.1%

Wash solution II

Tris.Cl, pH 7.5	10 mM
EDTA	2 mM
NaCl	1 M
BSA	0.1%
Triton X-100	0.1%

Western blot staining solution

Solution A:

Methanol	20 ml
4-Chloro-1-napthol	60 mg

Solution B:

Tris.Cl, pH 7.5 (10 mM)	80 ml
H_2O_2	100 µl

Solution A and B mixed immediately prior to use.

2xYT broth

Tryptone	1.6% w/v
Yeast extract	1.0% w/v
NaCl	0.5% w/v

APPENDIX 4 - MOLECULAR SIZE STANDARDS

DNA

λ *Hind*III digest

23.13 kbp
9.416 kbp
6.557 kbp
4.361 kbp
2.322 kbp
2.027 kbp
0.564 kbp
0.125 kbp

100 bp Ladder

1.5 kbp
1.4 kbp
1.3 kbp
1.2 kbp
1.1 kbp
1.0 kbp
0.9 kbp
0.8 kbp
0.7 kbp
0.6 kbp
0.5 kbp
0.4 kbp
0.3 kbp
0.2 kbp
0.1 kbp

Protein

Pre-stained standards

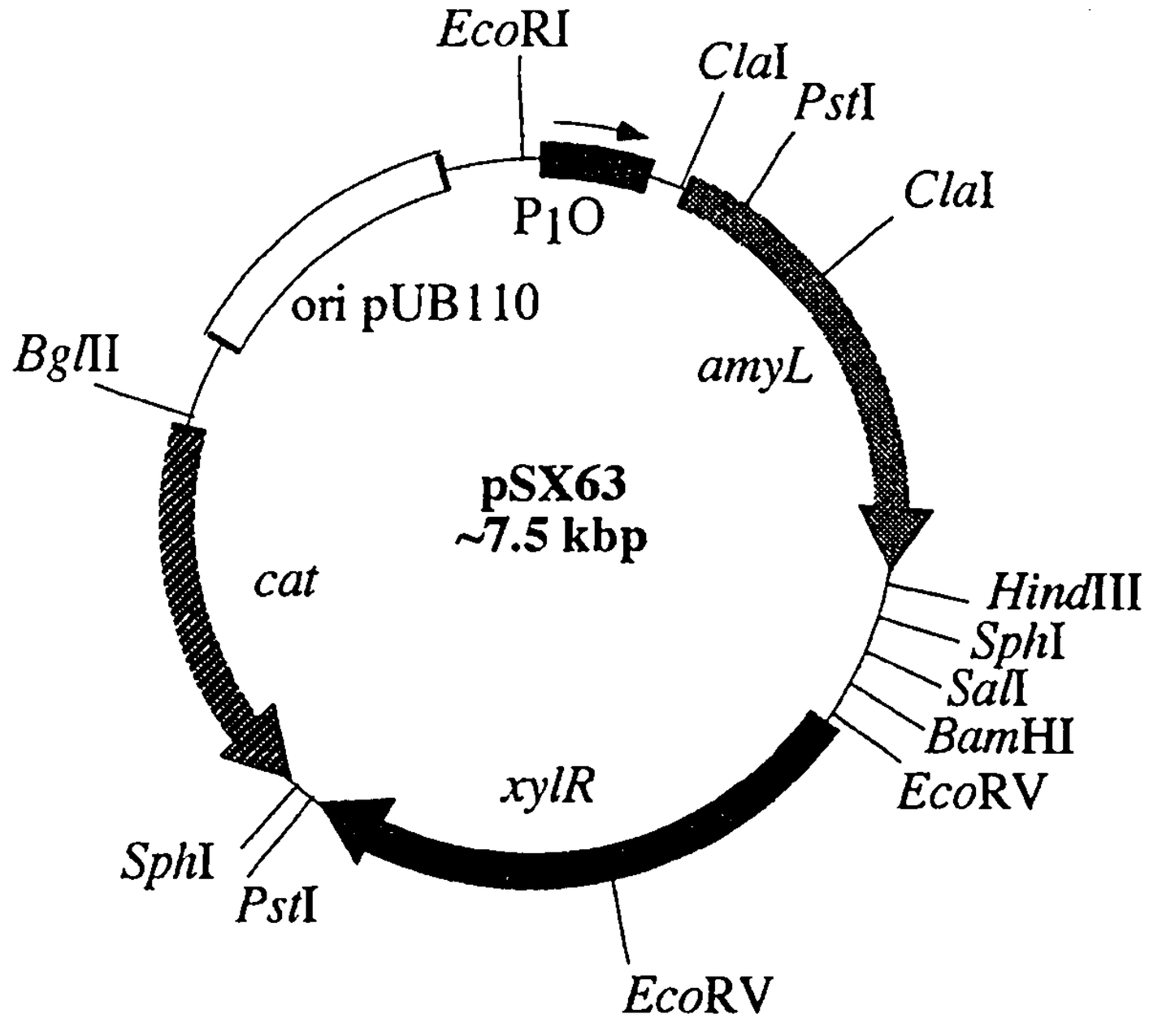
101 kDa
83 kDa
50.6 kDa
35.5 kDa
29.1 kDa
20.9 kDa

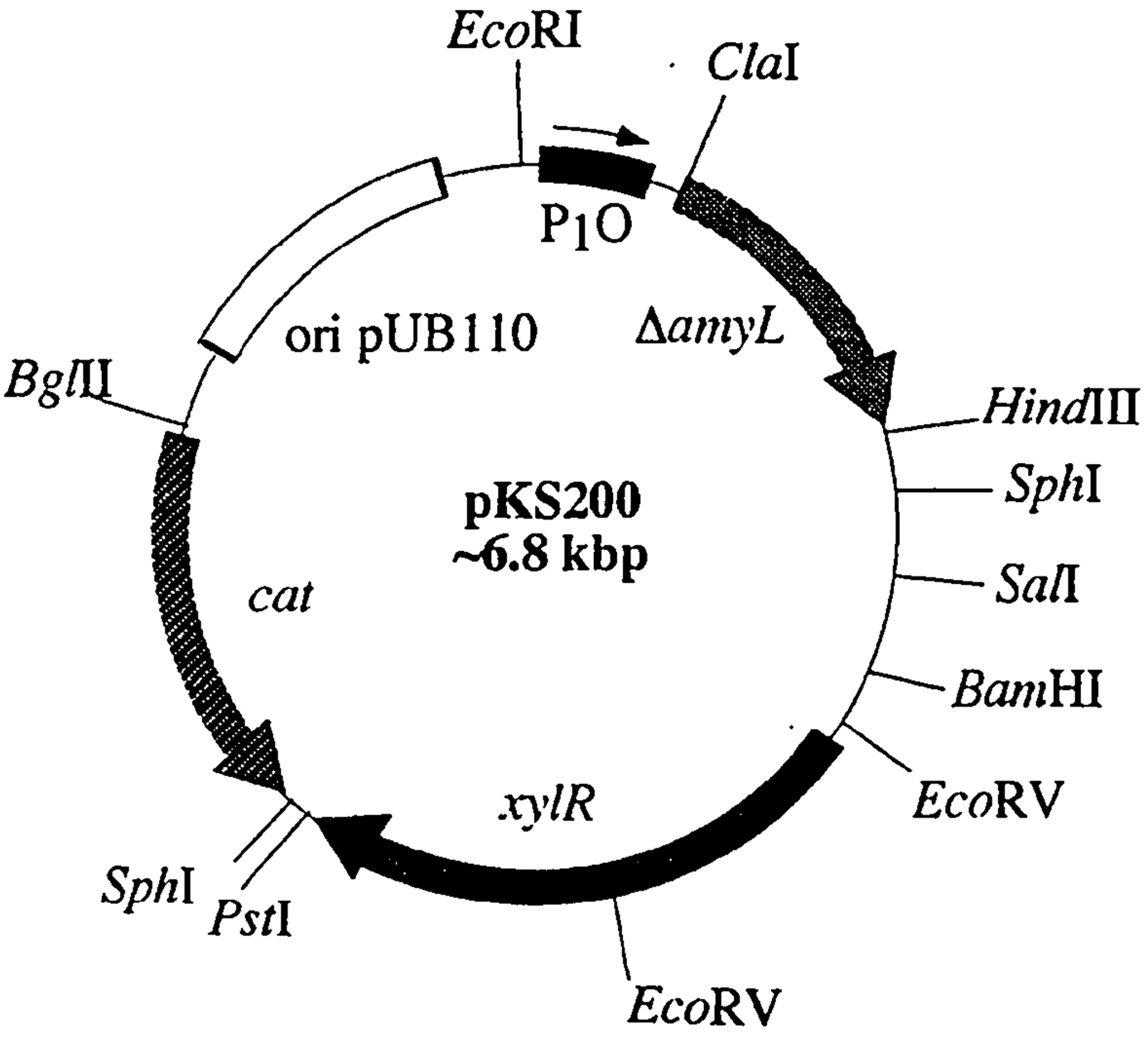
Rainbow ^{14}C methylated standards

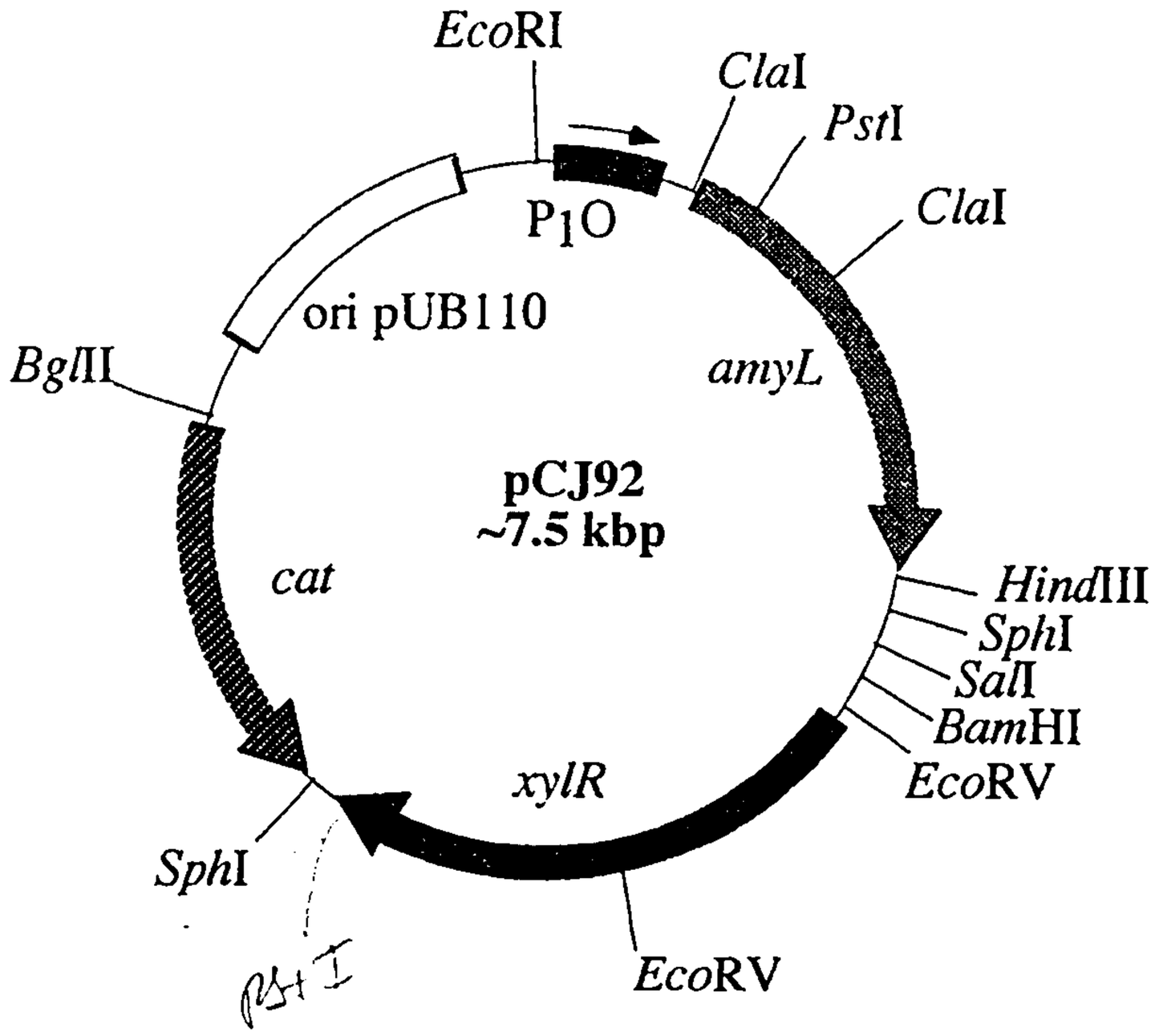
220 kDa
97.4 kDa
66 kDa
46 kDa
30 kDa
21.5 kDa
14.3 kDa

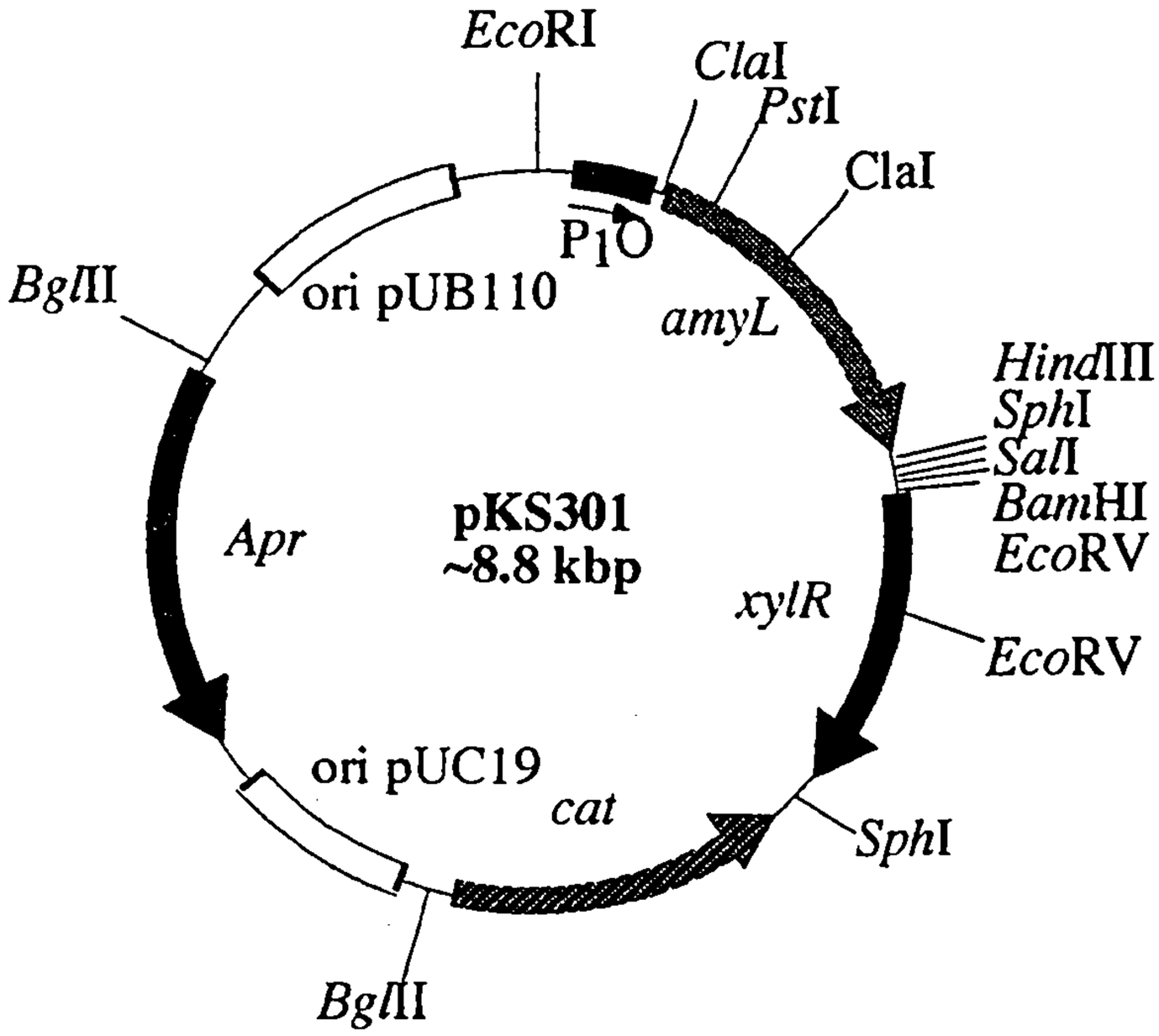
APPENDIX 5 - PLASMID MAPS

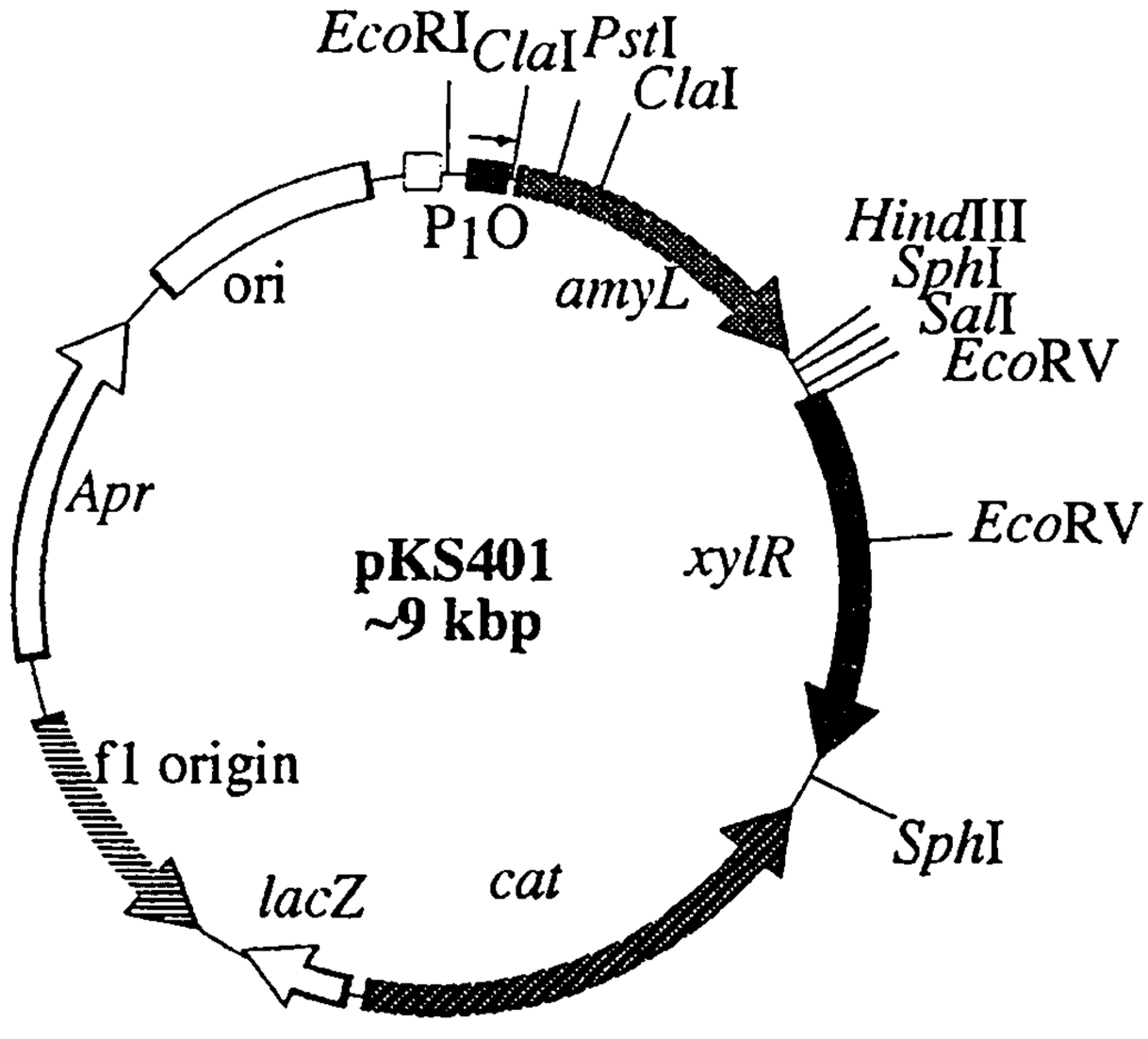
Plasmids not drawn to scale.

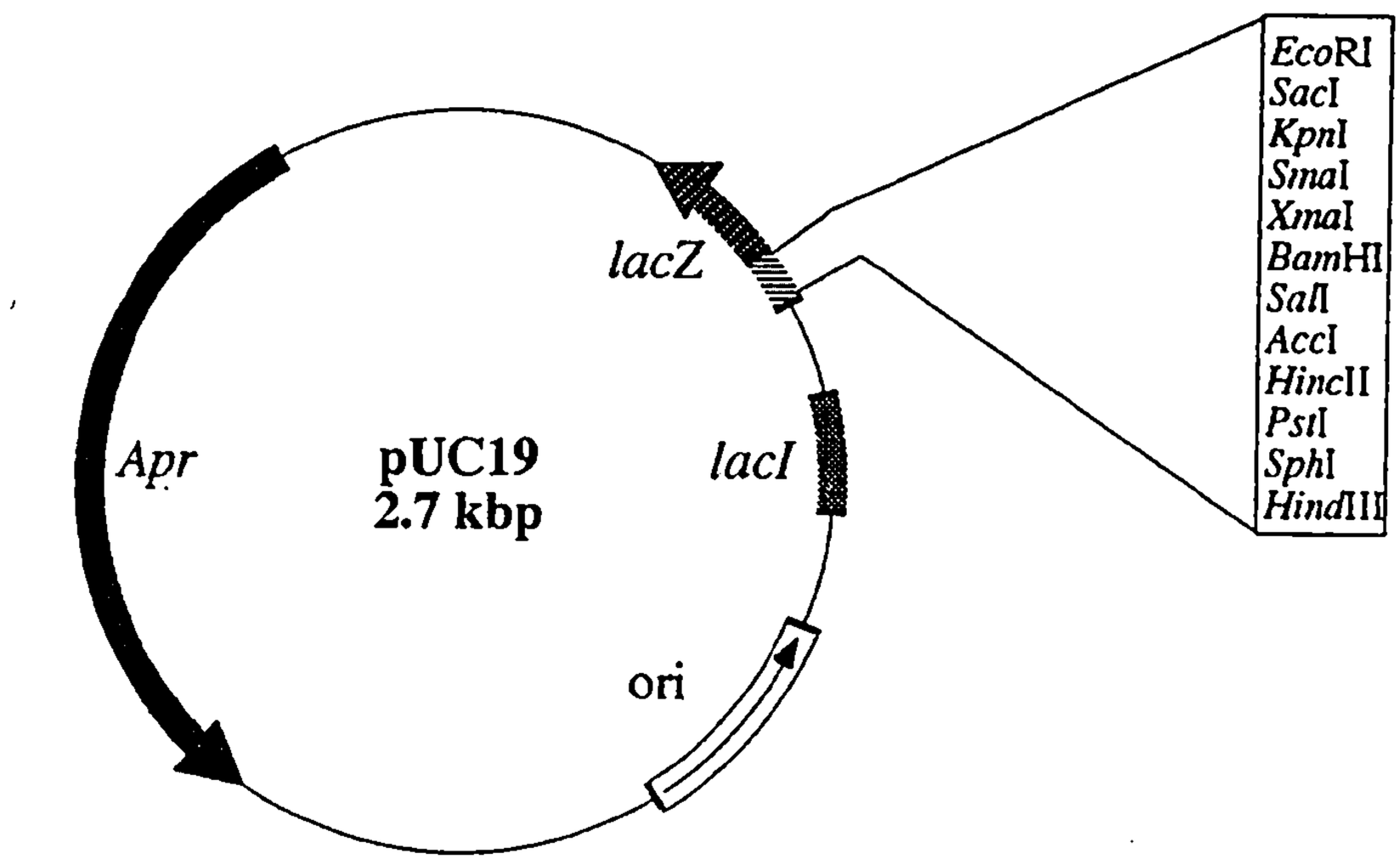


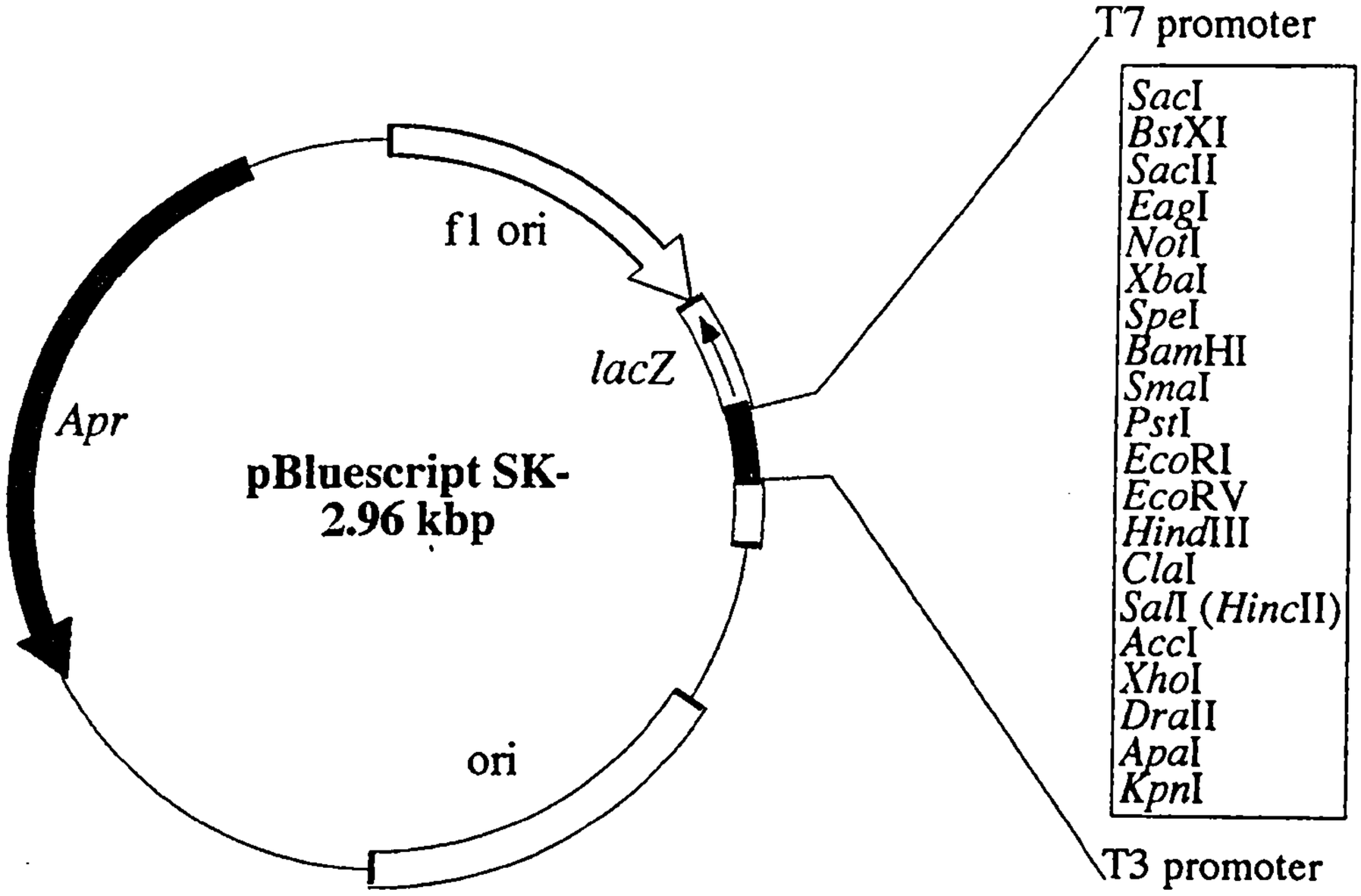



EcoRI
ClaI
PstI
ClaI
P1O
ori pUB110
BglII
amyL
pCJ92
~7.5 kbp
cat
HindIII
SphI
SalI
BamHI
EcoRV
xylR
SphI
EcoRV


EcoRI
ClaI
PstI
ClaI
P1O
BglII
ori pUB110
amyL
HindIII
SphI
SalI
BamHI
EcoRV
pKS301
~8.8 kbp
Apr
xylR
EcoRV
ori pUC19
cat
SphI
BglII


EcoRI ClaI PstI
ClaI
P1O
ori
amyL
HindIII
SphI
SalI
EcoRV
EcoRV
Apr
pKS401
~9 kbp
xylR
SphI
f1 origin
lacZ
cat


lacZ
Apr
pUC19
2.7 kbp
lacI
ori
EcoRI
SacI
KpnI
SmaI
XmaI
BamHI
SalI
AccI
HincII
PstI
SphI
HindIII


T7 promoter
SacI
BstXI
SacII
EagI
NotI
XbaI
SpeI
BamHI
SmaI
PstI
EcoRI
EcoRV
HindIII
ClaI
SalI (HincII)
AccI
XhoI
DraII
ApaI
KpnI
T3 promoter
f1 ori
lacZ
Apr
pBluescript SK-
2.96 kbp
ori